

THE USE OF OPTICAL ABERRATION COEFFICIENTS

by

P. W. FORD B. Sc (Hons.)

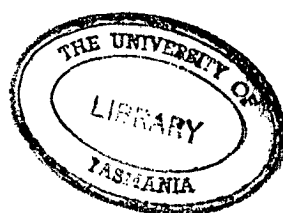
A Thesis submitted for examination for the degree of

DOCTOR OF PHILOSOPHY

UNIVERSITY OF TASMANIA

May, 1962.

Thesis
BIL
FORD



63/5733

Contents

Introduction	I
Part I	
SUMMARY OF THEORY	
1. Sign Conventions. Aberrations of a Ray	1
2. Paraxial Coefficients	4
3. Aberration Coefficients	8
4. "a" and "b" Aberration Coefficients	13
5. Calculation of Aberration Coefficients	21
6. Intrinsic Coefficients	30
Part II	
COMPUTER AND PROGRAMMES	
7. Computer Description	34
8. Aberration Coefficients Programme	40
9. Ray Trace Programme	45
10. Predicted Displacements from Aberration Coefficients	59
Part III	
NUMERICAL RESULTS	
11. Specifications of Systems Examined	63
12. Calculations Performed on the Optical Systems	69
13. Quality of Agreement between ε'_k and $\varepsilon'[2]$, $\varepsilon'[3]$	72
14. Criteria Indicating Accuracy of Predicted Displacements.	82

Contents

Part IV

APPLICATION OF ABERRATION COEFFICIENTS

15.	Use of Coefficients	86
16.	Derivatives of the Coefficients	93

APPENDIX 1

Final Aberration Coefficients of Tessar	96
---	----

APPENDIX 2

Portion of Biotar 12° Ray Trace Results	97
--	----

APPENDIX 3

Portion of Biotar 12° Predicted Displacements	98
--	----

APPENDIX 4

Formulae Used in Aberration Coefficients Programme	99
--	----

REFERENCES and PUBLISHED WORK.	104
--------------------------------	-----

INTRODUCTION.

An extensive theory of aberration coefficients of symmetrical optical systems has been developed by Buchdahl in his monograph "Optical Aberration Coefficients"¹ (hereafter called M) and extended in subsequent papers.^{2, 3, ...9.} The advantages resulting from the use of these coefficients rest in two important properties. Firstly, the one set of coefficients characterise systems of rays, that is, they apply simultaneously to all rays that traverse the optical system. Secondly the aberration coefficients are the sums of corresponding coefficients computed for each surface of the system (the contributions to the coefficients). This enables the action of the system on all rays to be analysed surface by surface and it is this that places a powerful tool in the hands of the designer.

Now, although there is only one set of coefficients for each system, it is an infinite set. Obviously the calculation of them all is impossible. So far, computing schemes have been designed for the computation of all the third, fifth and seventh order monochromatic coefficients⁴, the coefficients of ninth⁵ and eleventh⁹ order spherical aberration, and several of the more important chromatic coefficients (M Chapter XIII). Naturally, the aberrations of a system are not completely described by only these coefficients.

The object of this thesis is to examine the effectiveness of the first three orders of the monochromatic coefficients in the description of the aberrations of optical systems. As well as enabling a detailed analysis of a system, the coefficients and their surface contributions are of considerable use in the differential correction of a system following the initial design. The effectiveness of the coefficients in this field is also examined here.

The work has been restricted to monochromatic coefficients, since, after the initial design, the majority of design is carried

out in monochromatic light. Also, the systems studied here have been restricted to those containing spherical surfaces only. The extension of the work to aspherical surfaces is a matter of detail and not of method; as mentioned in M Section 55, the only change is that the "intrinsic" coefficients (Section 5, 6 of this thesis) contain additional terms which depend on the "extra-axial" curvatures of the aspheric surfaces. These additional terms in no way affect the general theory or the application of the coefficients. The aberration coefficients in no way indicate whether the system they represent is aspheric or not. However, in the construction of computing schemes for the coefficients, many simplifications can be introduced if only spherical surfaces are being considered, resulting in comparatively short schemes, e.g. compare M 81.3 with 84.23, 33, 44.

Since the publication of M, several misconceptions have occurred regarding the coefficients and the contributions to them by the surfaces. Some of these have been discussed in a paper by Cruickshank and Hills¹⁰. It is probable that these erroneous impressions have occurred as a result of 1) the multiplicity of symbols used in M, and 2) the iterative method used to derive the expressions for the coefficients.

About the symbolism little can be done. As mentioned in the preface to M "higher-order optics is a battle of symbolism, not of advanced mathematics." Consequently, where type allows, the symbolism in this thesis is that of M. The most notable exception is the use here of single and double underlines, the symbols so marked representing the bold-face type and the Gothic script of M respectively.

Iteration is quite familiar; everyone is acquainted with Newton's method for obtaining square roots, in which successively better approximations to the square root are obtained with each

iterative step. Perhaps as a result of this application, some workers have supposed that the aberration coefficients are approximate. This is entirely incorrect. As in the square root procedure, approximation is involved only by virtue of the fact that an infinite series has been terminated without regard to the remaining terms. The aberration coefficients are the coefficients of the terms in the infinite power series expansion for the "displacement" of a ray (See Section 1). When this series is terminated, we have only an approximation of the displacement. It is the accuracy of this approximation that is being examined here. However, the coefficients of the series are exact. After each iterative step explicit expressions for new higher-order coefficients are obtained, not better approximations of earlier ones.

Failure to realise this may have arisen as a result of the iteration in M being applied to a series containing more than one variable, which necessarily involves a large number of symbols. Therefore in Part I of this thesis two examples of iteration are presented, the second being more complex than the first. The second example introduces the idea of "intrinsic" coefficients, these being the basic coefficients from which all others are obtained. It is hoped that these examples will lead to a clearer understanding of the methods used in M.

By choosing suitable coordinates with which to specify a ray, considerable advantages ensue both in the computation and application of the aberration coefficients. These coordinate systems and the advantages thereof are also presented in Part I along with the basic theory of the coefficients.

The examination of the effectiveness of the coefficients in describing the aberrations of a number of representative systems entails considerable numerical work. Prior to this work, all

IV

calculations were performed on desk machines and it was obvious from the outset that this was too slow. For instance, a set of third, fifth and seventh order coefficients for a six surface system occupied two girls for four days. Thus my first task here was to learn to programme an electronic computer to perform this and other calculations. Accordingly, an approach was made to the University of New South Wales for the use of their machine, an English Electric "DEUCE". As a result of their very generous assistance, in about nine months I had written a programme for the computation of the first three orders of coefficients, which, incidentally, almost exceeded the 8,000 word capacity of the machine. It has subsequently been rewritten to achieve a reduction in computing time of about 20 percent. Using this programme, the coefficients for a six surface system can be obtained in about 5 minutes machine time.

Following this, two more programmes were written, one for general ray tracing and the other for the computation of the displacement of a ray using the aberration coefficients. Apart from a few special trigonometric ray traces, the entire numerical work in this thesis is based on the results of these three programmes.

Since these programmes were designed for general use in optical design, they have all been coded in basic machine language. The increase in time and expense for the programming is more than offset by the considerable economies in machine time realised during their subsequent extensive use.

The "DEUCE" is an interesting machine in that it uses mercury acoustic delay-lines for the high speed stores with a magnetic drum as the backing store. Notwithstanding this rather slow type of high speed storage the machine is quite fast in its operation. This is due mainly to the high digit frequency (1 megacycle), the arrangement of the arithmetic units and the considerable amount of information contained in one instruction word. The first section of Part II describes the design and operation of the computer. The remaining sections are devoted to the details of

the three programmes written for it.

In Part III the accuracy of the first three orders of aberration coefficients in describing the aberrations of optical systems is examined. Six modern photographic objectives have been used for this work. This is an extension of the work presented in a paper⁶ by Buchdahl in which he uses two systems as illustration, one of wide field and the other of large aperture. The quality of the predicted displacements has been judged on the appearance of tangential curves, annular curves and spot diagrams plotted for several pencils in each of the six systems.

In the tangential and annular curves are plotted the displacements of rays predicted from the first two and the first three orders of coefficients, as well as the true displacements obtained from ray traces. Predicted displacements calculated using the first three orders only are used in the spot diagram comparisons. The results of the comparisons have been tabulated in Part III (Table V).

It would be of advantage to the optical worker if some simple criterion could be found which would indicate the reliability of the coefficients in predicting displacements of rays. One such criterion, which I have termed an "angle criterion" was suggested in reference 6, "It is a matter of experience that reliance cannot be placed on the values of the predicted displacements if some $|\sin I|$ or $|\sin I'|$ (in a ray trace) exceed a value of about 0.8". This and other angle criteria have been examined in the light of the six additional systems used in Part III. Owing to the way the power series expansion of the displacement is developed, it appears that it is unlikely that any simple angle criterion will ever be completely satisfactory. It is my conclusion that an examination of certain annular curves afford the best means of assessing the reliability of the coefficients, the construction of these curves involving very little more work than that required for any of the angle criteria.

Part IV deals with the methods of application of the aberration coefficients in the analysis and differential correction of optical systems. These are the important applications; the mere construction of curves and spot diagrams can be done by ray tracing, which, however does not give any indication of why the system performs as it does, or how to improve it. On the other hand, the aberration coefficients and their derivatives, which characterise systems of rays, can indicate which surfaces require modification and as well can indicate with considerable accuracy the effect of simultaneous changes of parameters at several surfaces. When it is known that the coefficients adequately describe the aberrations of a system, their use is quite straight forward and is described in the paper by Cruickshank and Hills¹⁰. This is summarised in Part IV.

However, the main work in this part is concerned with a method that can be used when the first three orders of coefficients do not in themselves adequately describe the system. The method is based on a property of the coefficients which arises from the processes used in initial design. As a result of this property, it is still possible to analyse a system and to predict the effect of changes of parameters even when the coefficients are known not to adequately represent the aberrations. The method is illustrated using two of the systems already described in Part III.

In order to use the coefficients to predict the effect of changes of parameters, it is necessary to know at least the first derivatives of the coefficients with respect to the parameters. As yet, only the derivatives of the third order coefficients can be obtained directly from the expressions in M. While the necessary theory has been presented from which the derivatives of higher order coefficients may be found, the magnitude of the work involved has made this uneconomic at present. Should this ever be done, then the explicit expressions so obtained would enable by far the most satisfactory way of calculating the required derivatives. In fact,

if second derivatives are required then the theory presented in M will have to be developed. However, a knowledge of the first derivatives of the third, fifth and seventh order coefficients enables the effect of 5 percent changes to be predicted satisfactorily. These derivatives can be obtained by numerical means using the aberration coefficients programme. The accuracy of this method is discussed and illustrated with use of a cemented doublet. With the aid of certain identities between the derivatives and coefficients, developed from the theory in M, it is shown that the first derivatives can be computed quite accurately.

Acknowledgements.

I wish to thank my supervisor, Dr. F. D. Cruickshank, for his interest in the progress of this work. As an experienced optical designer, his discussions on the problems confronting optical workers were very helpful in indicating the type of work to be undertaken in this thesis. It was also under his auspices that the many trips to use the computer in Sydney were made possible.

The numerous discussions with Dr. H. A. Buchdahl concerning the theoretical and practical aspects of the aberration coefficients were invaluable. I cannot speak too highly of his assistance in this work, without which my task would have been much more difficult.

I am very grateful to the staff in charge of the "DEUCE" computer at the University of New South Wales, Sydney, for the way in which they placed their facilities at my disposal. Several times they worked overtime and rearranged time schedules to accommodate me, often at rather short notice. It was a pleasure to work with them.

I am indebted to Mrs. B. Brown who carried out the auxillary calculations used in this work. She also assisted in checking the coding and card-punching of the various programmes, as well as computing the results necessary to check their initial operation.

My thanks go to two typists, Mrs. P. Top and Mrs. G. Harman who cheerfully undertook the often complex typing of the manuscript.

Finally, I would like to acknowledge the generous assistance of the staff of the Photographic Department who carried out the printing of this thesis.

PART I - Summary of Theory

1. Sign Conventions. Aberrations of a Ray.

In M and this thesis, the following conventions for symbols and subscripts are used. Upper case letters refer to quantities associated with finite rays before refraction at a surface, the surfaces themselves being specified by subscripts 1, 2, ... j, ... k. The subscript k always refers to the last surface. Primed (') upper case letters denote the corresponding quantities after refraction. Lower case letters refer to quantities associated with rays lying everywhere in the infinitesimal neighbourhood of the axis (paraxial rays), that is, to quantities determined only by the laws of paraxial optics. All superscript and subscript conventions apply similarly to paraxial and non-paraxial symbols. Unless otherwise stated, symbols from which the surface subscript is omitted refer to quantities at any surface. These are the principal conventions which apply throughout, but additional superscripts and subscripts will be introduced as required.

Associated with each surface of the system is a left-handed set of rectangular cartesian axes with the origin at the pole of the surface. The x-axes of these coordinate systems lie along the axis of symmetry of the optical system, the positive direction being that in which light proceeds through the system. All the y-axes lie in the meridional plane. Let O_0 be the axial point of a plane, F_1 , in the object space, normal to the axis. Then rays from O_0 [l_{01} , 0, 0] whose paths through the system are determined only by the laws of paraxial optics, will intersect the axis in the final image space at a point O_k' [l_{k1} , 0, 0]. O_k' is defined as the axial point of the ideal image plane F_k' , conjugate to F_1 . Select any point O_1 in the object plane F_1 and denote its coordinates by [l_{01} , $-H_{y1}$, $-H_{z1}$], (Fig.1). By definition, if the optical system were perfect, all rays from O_1 which pass through the system would intersect in the image plane

2.

F_k in some point J whose coordinates $[l_{ok}, -h_{yk}, -h_{zk}]$ would be such that

$$h_{yk} = m_k' H_{y1}, \quad h_{zk} = m_k' H_{z1}$$

where m_k' is a constant of the system independent of H_{y1}, H_{z1} , called the paraxial magnification for the conjugate planes F_1, F_k' . J is called the ideal image point conjugate to O_1 ; and h_{yk}, h_{zk} the ideal image heights. In actual practice, however, a ray from O_1 will not, in general, pass through J , but through some neighbouring point $O_k' [l_{ok}', -H_{yk}', -H_{zk}']$. The aberration $\varepsilon_{yk}, \varepsilon_{zk}$ (Fig.2) of the ray is defined as

$$\varepsilon_{yk} = H_{yk} - h_{yk}, \quad \varepsilon_{zk} = H_{zk} - h_{zk}.$$

To simplify the writing, all symbols singly underlined shall be taken as referring to both the y- and z- components of the quantity in question. Hence the preceding equation will be written

$$\underline{\varepsilon}_k = \underline{H}_k - \underline{h}_k. \quad (1.1)$$

Since each refracting surface has an object and image plane conjugate to F_1 , the j th image plane being also the $(j+1)$ th object plane, the aberrations of a ray can be defined at these planes in a manner similar to (1.1). Thus, at the j th object plane,

$$\underline{\varepsilon}_j = \underline{H}_j - \underline{h}_j, \quad (1.2)$$

and at the j th image plane

$$\underline{\varepsilon}_j' = \underline{H}_j' - \underline{h}_j' \quad (1.3)$$

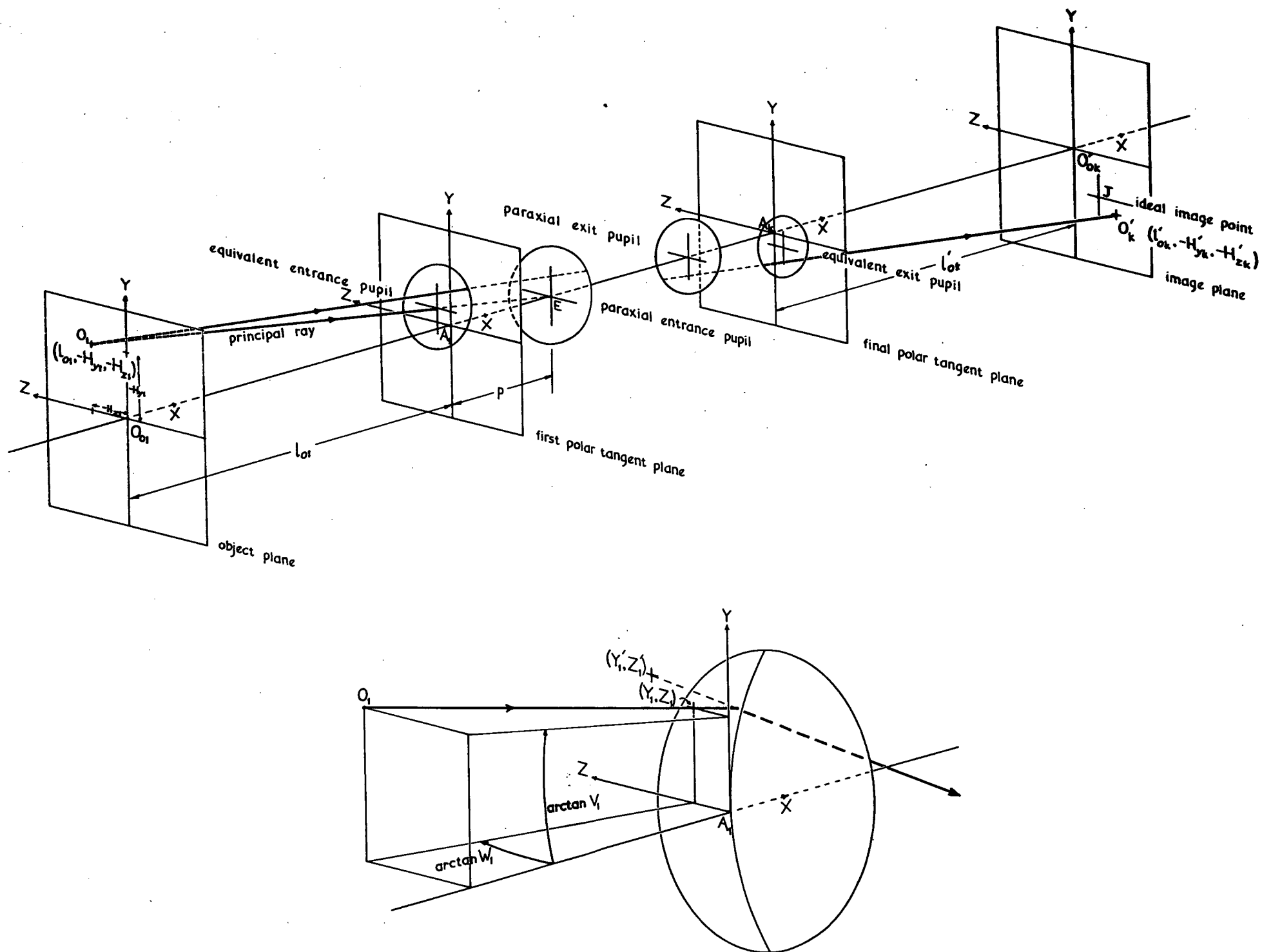


Fig. 1 Coordinate system and canonical coordinates

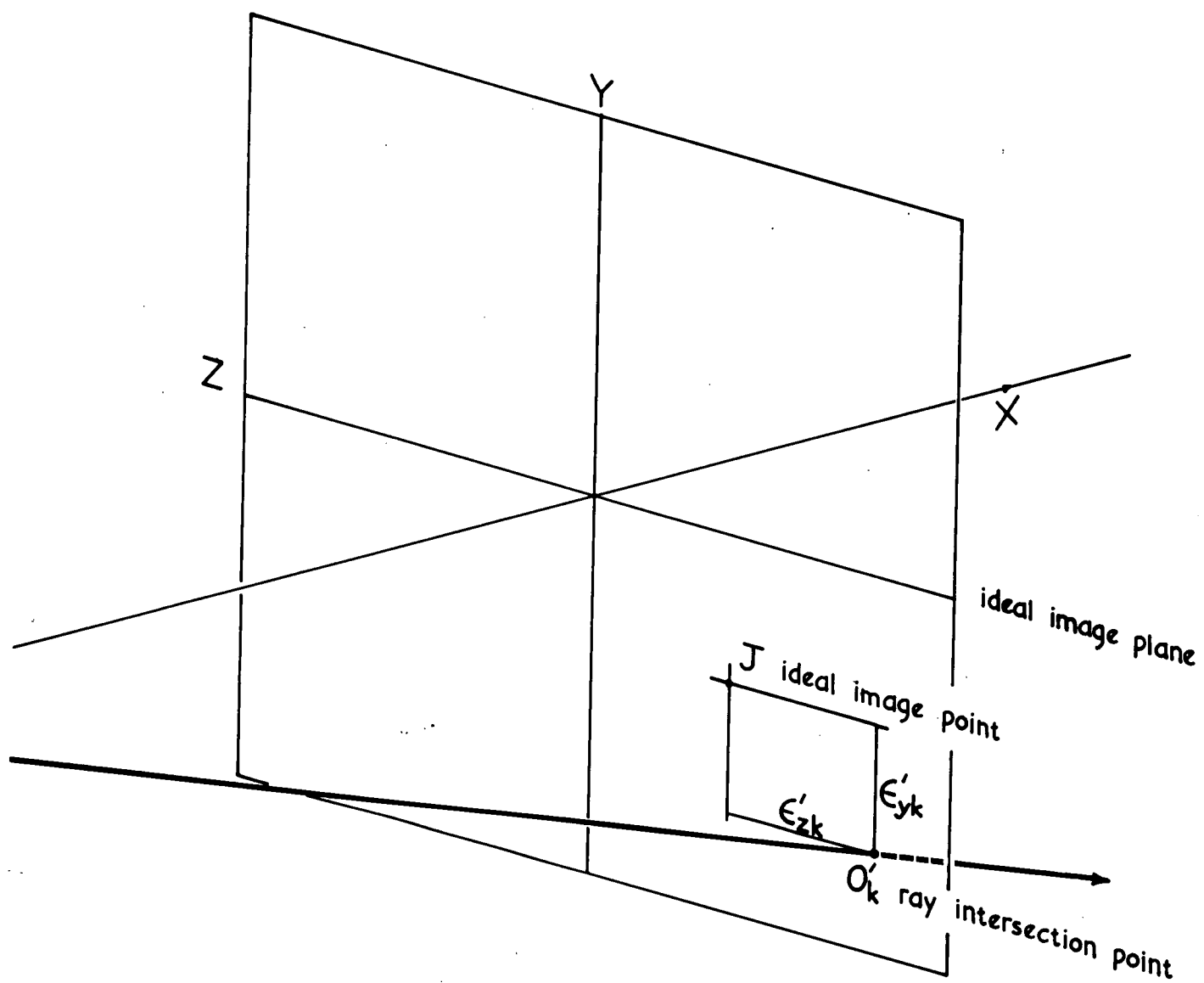


Fig. 2 Aberrations (or displacements) ϵ'_{yk} , ϵ'_{zk}

A ray incident at any surface will be specified by the four canonical variables Y, Z, V, W . Y and Z are cartesian coordinates of the point of intersection of the ray with the plane tangent to the refracting surface at its pole while V and W are related to the direction cosines (α, β, γ) of the ray by the expressions

$$V = -\beta/\alpha, \quad W = -\gamma/\alpha \quad (1.4)$$

whence it follows that

$$\alpha = [1 + V^2 + W^2]^{-\frac{1}{2}} \quad (1.5)$$

In accordance with the convention just introduced \underline{Y} is to be interpreted as standing for both Y and Z , and \underline{V} as standing for both V and W .

If $\bar{x}, \bar{y}, \bar{z}$ are the coordinates of a current point on a ray, then the equation of the ray may be written

$$\bar{y} = \underline{Y} - \underline{V} \bar{x} \quad (1.6)$$

If we put $\bar{x} = d_j'$, the axial distance between surfaces j and $j+1$, then \bar{y} will be the coordinates of the point of intersection of the ray with the $(j+1)$ th polar tangent plane.

Thus

$$\underline{Y}_{j+1} = \underline{Y}_j' - d_j' \underline{V}_j'; \quad \underline{V}_{j+1} = \underline{V}_j'.$$

It is convenient to omit the subscript j and to replace the subscript $j+1$ simply by the subscript $+$. Then we have

$$\underline{Y}_+ = \underline{Y}' - d' \underline{V}'; \quad \underline{V}_+ = \underline{V}'. \quad (1.7)$$

4.

Equations (1.7) are the transfer equations for canonical variables.

2. Paraxial Coefficients

Defining the symbol Δ to be the change in a quantity on refraction at a surface e.g. $\Delta X = X' - X$, it is shown in M Section 4 that

$$\underline{I} = c\underline{Y} - \underline{V} \quad (2.1)$$

$$\Delta(N \propto \underline{I}) = 0 \quad (2.2)$$

$$\text{and also} \quad \Delta\underline{Y} = x\Delta\underline{V} \quad (2.3)$$

where $c (=1/r)$ is the curvature of the surface. For paraxial rays these relations reduce to

$$\left. \begin{aligned} \Delta N \underline{I} &= 0 \\ c\Delta\underline{Y} &= \Delta(\underline{I} + \underline{V}) = 0 \end{aligned} \right\} \quad (2.4)$$

since $\alpha, \alpha' \rightarrow 1$, $x\Delta\underline{V} \rightarrow 0$.

Using (2.4) and (1.7) we get the paraxial recurrence relations

$$\left. \begin{aligned} \underline{y}_{j+1} &= [1 - (1-k)cd']\underline{y}_j - kd'\underline{v}_j \\ \underline{v}_{j+1} &= (1-k)c\underline{y}_j + k\underline{v}_j \end{aligned} \right\} \quad (2.5)$$

where $k = N/N'$.

Thus we can obtain \underline{y}_{j+1} , \underline{v}_{j+1} from \underline{y}_j , \underline{v}_j . Since \underline{y}_{j+1} , \underline{v}_{j+1} are linearly related to \underline{y}_j , \underline{v}_j , then \underline{y}_j , \underline{v}_j are linearly related to \underline{y}_1 , \underline{v}_1 and the constants involved depend only on the constitution of the system. Hence

$$\left. \begin{aligned} \underline{y}_j &= y_{pj} \underline{y}_1 + y_{qj} \underline{v}_1 \\ \underline{v}_j &= v_{pj} \underline{y}_1 + v_{qj} \underline{v}_1 \end{aligned} \right\} \quad (2.6)$$

the y_{pj} , y_{qj} , v_{pj} , v_{qj} being called the "paraxial coefficients" of the j th surface. Notice that the relations for y_j and z_j contain the same coefficients, and that the same is true of the expressions for v_j and w_j . Other paraxial coefficients can be formed from these by linear combination, e.g.

$$\left. \begin{aligned} i_{pj} &= c y_{pj} - v_{pj} \\ i_{qj} &= c y_{qj} - v_{qj} \end{aligned} \right\} \quad (2.7)$$

so that

$$\underline{i}_j = i_{pj} \underline{y}_1 + i_{qj} \underline{v}_1 \quad (2.8)$$

and these will be used as the need arises.

The paraxial "p" coefficients can be most easily obtained by tracing a ray by means of (2.5) whose formal starting data are $y_1 = 1$, $v_1 = 0$ (p-ray). Then the values of y , v at each surface are y_p , v_p . Similarly the "q" coefficients can be obtained by tracing a second ray having the starting data $y_1 = 0$, $v_1 = 1$ (q-ray).

An important identity between the paraxial coefficients is obtained as follows.

Consider two arbitrary paraxial rays, one of which is a tangential ray, denoted by a bar over the symbols. Defining a quantity $\underline{\lambda}$ by

$$\underline{\lambda} = N(\bar{y}v - \bar{v}y), \quad (2.9)$$

then it is shown in M Section 5(b) that $\underline{\lambda}$ is an optical invariant, that is, its value is constant throughout the system for the pair of rays considered. If, in (2.9), the tangential ray is one from the axial point of the object, and the second ray passes through the object point, then $\underline{\lambda}$ can be shown to be the

Lagrange invariant, i.e. $\lambda = N v_0 h$. Expressing the paraxial variables in (2.9) in terms of their values at the first surface by means of (2.6), one obtains

$$\lambda_j = (N_j/N_1) (y_{pj} v_{qj} - y_{qj} v_{pj}) \lambda_1.$$

Since $\lambda_1 = \lambda_j$,

$$y_{pj} v_{qj} - y_{qj} v_{pj} = N_1/N_j. \quad (2.10)$$

Apart from enabling the computation of the paraxial coefficients to be checked, this identity (2.10) is very useful in the simplification of many expressions which occur in the development of the aberration coefficient theory. Several other identities can be obtained from it, for example

$$v_p v_q' - v_q v_p' = c N_1 (k - 1)/N \quad (2.11)$$

$$y_p v_q' - y_q v_p' = N_1/N^2 \quad (2.12)$$

$$y_p i_q - y_q i_p = -N_1/N. \quad (2.13)$$

To simplify the notation, we introduce the convention that if a prime is attached to the left of a symbol, then the symbol is divided by N_1 . Such a prime will be termed an ante-prime. Thus (2.10) becomes

$$y_p v_q - y_q v_p = 1/N. \quad (2.14)$$

The identity (2.14) is used in the following derivation for the focal lengths of an optical system.

Let the axial points of the front and rear principal planes be P_0 and P_0' respectively, these planes, P, P' , having the property that the paraxial magnification associated with them

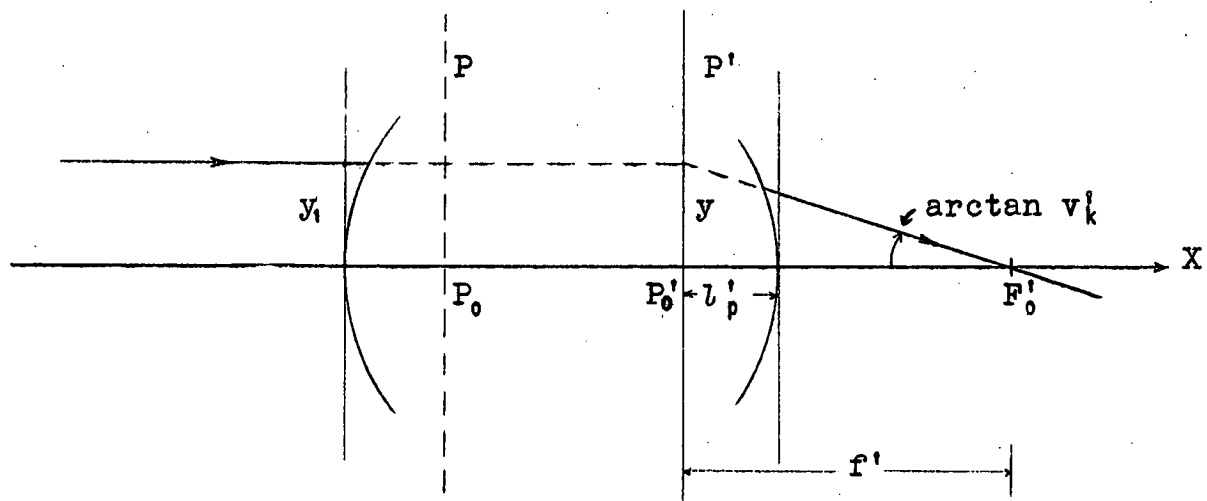


Fig. 3

Posterior focal length

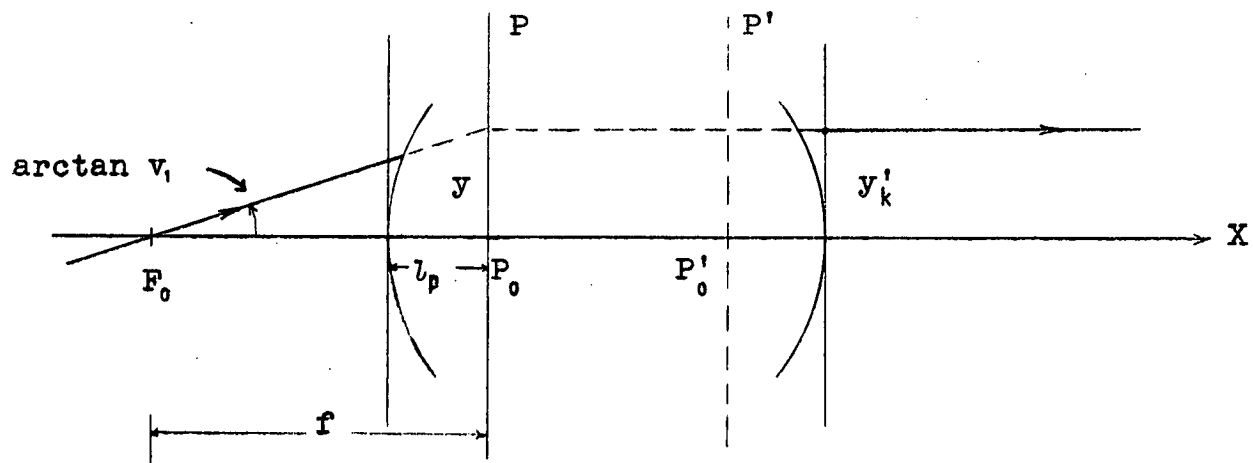


Fig. 4

Anterior focal length

is unity. Let F'_0 be the axial point of intersection in the image space of a paraxial ray incident on the system parallel with the axis. Then the posterior focal length, f' , of the system is defined as $f' = P'_0 F'_0$, (Fig.3). Likewise, if F_0 is a point on the axis in the object space such that paraxial rays from it are parallel to the axis in the image space, then the anterior focal length, f , is defined as $f = P_0 F_0$, (Fig.4). Note that the lengths are measured from the principal planes.

As can be seen in Fig. 3

$$\begin{aligned} f' &= y/v'_k \\ &= y_1/v'_k . \end{aligned}$$

By (2.6)

$$\begin{aligned} v'_k &= v_{pk} y_1 + v_{qk} v_1 \\ &= v_{pk} y_1 , \text{ since } v_1 = 0 . \end{aligned}$$

Thus

$$f' = 1/v_{pk} . \quad (2.15)$$

For the anterior focal length, (Fig.4)

$$f = y/v_1 = y'_k/v_1 ,$$

and

$$y'_k = y_{pk} y_1 + y_{qk} v_1 , \quad (2.16)$$

$$v'_k = v_{pk} y_1 + v_{qk} v_1 \equiv 0 . \quad (2.17)$$

Thus

$$y_1 = - v_{qk} v_1 / v_{pk}$$

8.

and (2.16) becomes

$$\begin{aligned} y'_k &= v_1 (y'_{qk} v'_{pk} - y'_{pk} v'_{qk}) / v'_{pk} \\ &= -N_1 v_1 / N'_k v'_{pk} \quad \text{from (2.10).} \end{aligned}$$

Hence

$$f = -N_1 / N'_k v'_{pk} = -\frac{N_1}{N'_k} f' \quad (2.18)$$

The distance l_p of the first principal plane from the first surface is given by

$$l_p = (N_1 / N'_k - v'_{qk}) / v'_{pk} \quad (2.19)$$

l'_p , the distance of the rear principal plane from the last surface is given by

$$l'_p = (y'_{pk} - 1) / v'_{pk} \quad (2.20)$$

3. Aberration Coefficients

(a) In section 1, the aberration of a ray in the j th image plane, $\underline{\epsilon}'$, was defined as

$$\underline{\epsilon}' = \underline{H}' - \underline{h}' \quad (3.1)$$

$$\text{Thus } \underline{H}' = \underline{h}' + \underline{\epsilon}' \quad (3.2)$$

Multiplying throughout by $N'v'_0$, (the subscript "o" in this context denoting quantities associated with rays from the axial point of the object) we get

$$N'v'_0 \underline{H}' = N'v'_0 \underline{h}' + N'v'_0 \underline{\epsilon}' \quad (3.3)$$

If l_j is the axial distance of the j th image plane from the j th surface, from (1.6) we get

$$\underline{h}' = l_j \underline{v}' - \underline{y}' \quad (3.4)$$

remembering the sign convention for \underline{h}' .

Thus

$$\begin{aligned} N' v_j \underline{h}' &= N' v_j (l_j \underline{v}' - \underline{y}') \\ &= N' v_j [(y_j / v_j) \underline{v}' - \underline{y}'] \\ &= N' (y_j \underline{v}' - v_j \underline{y}') \end{aligned}$$

which is the invariant $\underline{\lambda}$, by (2.9). Hence (3.3) becomes

$$N' v_j \underline{H}' = \underline{\lambda} + N' v_j \underline{\varepsilon}' \quad (3.5)$$

Writing $\underline{\Delta}$ for $N v_0 \underline{H}$, and $\underline{\hat{\varepsilon}}$ for $N v_0 \underline{\varepsilon}$,

$$\underline{\Delta \Delta} = \underline{\Delta \hat{\varepsilon}} \quad (3.6)$$

Incidentally, since $N' v_0 \underline{h}'$ is an invariant

$$N' v_0 \underline{h}' = N_1 v_{01} \underline{h}_1 \quad (3.7)$$

or $\underline{h}' = m \underline{h}_1$, where $m = N_1 v_{01} / N' v_0$,

and is termed the paraxial magnification associated with the planes F_1 and F_j . The paraxial magnification associated with the final ideal image plane F_k is

$$m_k' = N_1 v_{01} / N_k' v_{0k} \quad (3.8)$$

Now, the j th ideal image plane is also the $(j+1)$ th object plane, hence

$$\underline{\varepsilon}_j = \underline{\varepsilon}_{j+1}. \quad (3.9)$$

Also $N_j = N_{j+1}$ and $v_{0j} = v_{0(j+1)}$, thus

$$\hat{\underline{\varepsilon}}_j = \hat{\underline{\varepsilon}}_{j+1}. \quad (3.10)$$

Now

$$\begin{aligned} \hat{\underline{\varepsilon}}_k &= \hat{\underline{\varepsilon}}_1 - \hat{\underline{\varepsilon}}_1 + \hat{\underline{\varepsilon}}_2 - \hat{\underline{\varepsilon}}_2 + \dots - \hat{\underline{\varepsilon}}_k + \hat{\underline{\varepsilon}}_k' \\ &= \hat{\underline{\varepsilon}}_1 - \hat{\underline{\varepsilon}}_1 + \hat{\underline{\varepsilon}}_1' - \hat{\underline{\varepsilon}}_2 + \dots - \hat{\underline{\varepsilon}}_k + \hat{\underline{\varepsilon}}_k' \quad \text{from (3.10)} \\ &= \hat{\underline{\varepsilon}}_1 + (\hat{\underline{\varepsilon}}_1' - \hat{\underline{\varepsilon}}_1) + (\quad) \dots + (\hat{\underline{\varepsilon}}_k' - \hat{\underline{\varepsilon}}_k) \\ &= \hat{\underline{\varepsilon}}_1 + \Delta\hat{\underline{\varepsilon}}_1 + \Delta\hat{\underline{\varepsilon}}_2 + \dots + \Delta\hat{\underline{\varepsilon}}_k. \end{aligned} \quad (3.11)$$

However, $\hat{\underline{\varepsilon}}_1 \equiv 0$, thus, using (3.6)

$$\hat{\underline{\varepsilon}}_k' = \sum_{i=1}^k \Delta\hat{\underline{\varepsilon}}_i. \quad (3.12)$$

$$\Delta\hat{\underline{\varepsilon}} = N'v_0'H' - Nv_0H$$

and

$$\underline{H} = l_0 \underline{V} - \underline{Y}, \quad \text{from (1.4), thus}$$

$$\Delta\hat{\underline{\varepsilon}} = \Delta[N(v_0 \underline{Y} - y_0 \underline{V})]. \quad (3.13)$$

Since $\underline{Y}, \underline{V}$, depend on $\underline{Y}_1, \underline{V}_1$, $\Delta\hat{\underline{\varepsilon}}$ may be expanded as a power series in $\underline{Y}_1, \underline{V}_1$. Due to the rotational symmetry of the optical system, if $\underline{Y}_1, \underline{V}_1$ are reversed in sign, $\underline{\varepsilon}_k'$ reverses its sign, and therefore also must $\Delta\hat{\underline{\varepsilon}}$. Thus $\Delta\hat{\underline{\varepsilon}}$ must be a sequence of polynomials of odd degree in $\underline{Y}_1, \underline{V}_1$. Furthermore, $\Delta\hat{\underline{\varepsilon}}$ tends to zero as $\underline{Y}_1, \underline{V}_1$ become infinitesimal, that is, in the paraxial

limit, so that the series for $\Delta\Lambda$ does not contain linear terms. Thus the expansion for $\Delta\Lambda$ can be written in the form

$$\Delta\Lambda = \sum_{n=1}^{\infty} \sum_{\mu=0}^n \sum_{\nu=0}^{\mu} (g_{\mu\nu}^{(n)} \underline{Y}_1 + \bar{g}_{\mu\nu}^{(n)} \underline{V}_1) \xi_1^{n-\mu} \eta_1^{\mu-\nu} \zeta_1^{\nu} \quad (3.14)$$

where

$$\begin{aligned} \xi_1 &= Y_1^2 + Z_1^2, \\ \eta_1 &= Y_1 V_1 + Z_1 W_1, \\ \zeta_1 &= V_1^2 + W_1^2. \end{aligned} \quad (3.15)$$

For convenience, the early terms of the series (3.14) are written

$$\begin{aligned} \Delta\Lambda &= (a\underline{Y}_1 \xi_1 + \bar{a}\underline{V}_1 \xi_1 + b\underline{Y}_1 \eta_1 + \bar{b}\underline{V}_1 \eta_1 + c\underline{Y}_1 \zeta_1 + \bar{c}\underline{V}_1 \zeta_1) + \\ & (s_1 \underline{Y}_1 \xi_1^2 + \bar{s}_1 \underline{V}_1 \xi_1^2 + s_2 \underline{Y}_1 \xi_1 \eta_1 + \bar{s}_2 \underline{V}_1 \xi_1 \eta_1 + \\ & s_3 \underline{Y}_1 \xi_1 \zeta_1 + \bar{s}_3 \underline{V}_1 \xi_1 \zeta_1 + s_4 \underline{Y}_1 \eta_1^2 + \bar{s}_4 \underline{V}_1 \eta_1^2 + s_5 \underline{Y}_1 \eta_1 \zeta_1 + \\ & \bar{s}_5 \underline{V}_1 \eta_1 \zeta_1 + s_6 \underline{Y}_1 \zeta_1^2 + \bar{s}_6 \underline{V}_1 \zeta_1^2) + \\ & (t_1 \underline{Y}_1 \xi_1^3 + \dots + \bar{t}_{10} \underline{V}_1 \zeta_1^3) + O(9), \end{aligned}$$

where $O(9)$ denotes terms of degree not less than 9 in the coordinates.

Writing

$$G_{\mu\nu j}^{(n)} = \sum_{i=1}^{j-1} g_{\mu\nu i}^{(n)} \quad (3.16)$$

$$\text{and} \quad G_{\mu\nu j}^{(n)'} = \sum_{i=1}^j g_{\mu\nu i}^{(n)'}$$

and similarly for the barred coefficients, (3.12) becomes

$$\hat{\epsilon}_k = \sum_{n=1}^{\infty} \sum_{\mu=0}^n \sum_{\nu=0}^{\mu} (G_{\mu\nu k}^{(n)'} \underline{Y}_1 + \bar{G}_{\mu\nu k}^{(n)'} \underline{V}_1) \xi_1^{n-\mu} \eta_1^{\mu-\nu} \zeta_1^{\nu}. \quad (3.17)$$

Adopting the conventions in (3.16), (3.17) can be written

$$\hat{\varepsilon}k = (Ak \underline{Y}_1 \xi_1 + \dots + \bar{C}k \underline{V}_1 \eta_1) + (S'k \underline{Y}_1 \xi_1^2 + \dots + \bar{S}k \underline{V}_1 \eta_1^2) + (Tk \underline{Y}_1 \xi_1^3 + \dots + \bar{T}k \underline{V}_1 \eta_1^3) + O(9). \quad (3.18)$$

The coefficients $G_{\mu\nu k}^{(n)'}$, $\bar{G}_{\mu\nu k}^{(n)'}$ are called the (augmented) canonical aberration coefficients of order n of the system, and the $g_{\mu\nu j}^{(n)}$, $\bar{g}_{\mu\nu j}^{(n)}$ are termed the contributions to these by the j th surface. $G_{\mu\nu j}^{(n)}$, $\bar{G}_{\mu\nu j}^{(n)}$ are called the intermediate canonical aberration coefficients. Note that $G_{\mu\nu 1}^{(n)} = \bar{G}_{\mu\nu 1}^{(n)} \equiv 0$.

(b) $\Delta\Lambda_j$ depends linearly on y_{01} , v_{01} , by virtue of (3.13), (2.6). Thus the coefficients in (3.14) must depend linearly on y_{01} , v_{01} , so that

$$\begin{aligned} g_{\mu\nu}^{(n)} &= g_{\mu\nu p}^{(n)} y_{01} + g_{\mu\nu q}^{(n)} v_{01}, \\ \bar{g}_{\mu\nu}^{(n)} &= \bar{g}_{\mu\nu p}^{(n)} y_{01} + \bar{g}_{\mu\nu q}^{(n)} v_{01}. \end{aligned} \quad (3.19)$$

Then the "p" aberration coefficients are

$$G_{\mu\nu pk}^{(n)'} = \sum_{i=1}^k g_{\mu\nu pi}^{(n)}$$

and the "q" aberration coefficients are

$$G_{\mu\nu qk}^{(n)'} = \sum_{i=1}^k g_{\mu\nu qi}^{(n)},$$

and similarly for the barred coefficients.

Thus $\hat{\varepsilon}k$ can be written as the sum of two series, one containing only "p" coefficients, and the other containing the "q" coefficients, in the following manner.

$$\begin{aligned} \hat{\varepsilon}k &= \sum \sum \sum (G_{\mu\nu pk}^{(n)'} \underline{Y}_1 + \bar{G}_{\mu\nu pk}^{(n)'} \underline{V}_1) \xi_1^{\mu-\nu} \eta_1^{\mu-\nu} \zeta_1^{\nu} y_{01} \\ &+ \sum \sum \sum (G_{\mu\nu qk}^{(n)'} \underline{Y}_1 + \bar{G}_{\mu\nu qk}^{(n)'} \underline{V}_1) \xi_1^{\mu-\nu} \eta_1^{\mu-\nu} \zeta_1^{\nu} v_{01}. \end{aligned} \quad (3.20)$$

In conformity with (3.18),

$$\begin{aligned}
G_{00pk}^{(1)'} &= A_{pk}^i, & \bar{G}_{00pk}^{(1)'} &= \bar{A}_{pk}^i, & G_{00qk}^{(1)'} &= A_{qk}^i, \\
\bar{G}_{00qk}^{(1)'} &= \bar{A}_{qk}^i, & G_{10pk}^{(1)'} &= B_{pk}^i, & \dots &
\end{aligned} \tag{3.21}$$

The correspondence between the alphabetic symbols and the various values of μ, ν in the different orders is given in the following table.

Third order $n=1$	(μ, ν)	00 A	10 B	11 C						
Fifth order $n=2$		00 S ₁	10 S ₂	11 S ₃	20 S ₄	21 S ₅	22 S ₆			
Seventh order $n=3$		00 T ₁	10 T ₂	11 T ₃	20 T ₄	21 T ₅	22 T ₆	30 T ₇	31 T ₈	32 T ₉ 33 T ₁₀

4. "a" and "b" Aberration Coefficients

If the object tends to infinity, i.e. $l_{01} \rightarrow \infty$, then $v_{01} \rightarrow 0$ and (3.20) becomes

$$\hat{\varepsilon}_k^i = \sum \sum \sum (G_{\mu\nu pk}^{(n)'} \underline{y}_i + \bar{G}_{\mu\nu pk}^{(n)'} \underline{v}_i) \xi_1^{\mu-\nu} \eta_1^{\mu-\nu} \zeta_1^{\nu} y_{01} \quad (4.1)$$

Now

$$\begin{aligned}
\hat{\varepsilon}_k^i &= N_k^i v_{0k}^i \varepsilon_k^i \\
&= N_k^i \varepsilon_k^i (v_{pk}^i y_{01} + v_{qk}^i v_{01}) \\
&= N_k^i v_{pk}^i \varepsilon_k^i y_{01}
\end{aligned}$$

Thus, omitting the subscripts and primes of the coefficients,

$$\varepsilon_k^i = \mu [\sum \sum \sum (G_p^{(n)} \underline{y}_i + \bar{G}_p^{(n)} \underline{v}_i) \xi_1^{\mu-\nu} \eta_1^{\mu-\nu} \zeta_1^{\nu}], \quad (4.2)$$

where $\mu = 1/Nk'v_{pk}$. That is, only the p coefficients are required for the description of $\underline{\epsilon}_k$.

Consider a pencil of rays from an infinitely distant object point in the tangential plane. Let this pencil make an angle V_1 with the axis ($W_1 = 0$). Let p be the distance of the paraxial entrance pupil from the pole of the first surface (See M Section 33(a)). Then taking polar coordinates ρ, θ in the first polar tangent plane, with the intersection point of the principal ray as origin, Fig.5,

$$\begin{aligned} Y_1 &= pV_1 + \rho \cos \theta, & Z_1 &= \rho \sin \theta \\ V_1 &= V_1, & W_1 &= 0. \end{aligned}$$

Equation (4.2) becomes

$$\begin{aligned} \underline{\epsilon}_k &= \mu [\Sigma \Sigma \Sigma (G_{\mu\nu\rho k}^{(n)'} (pV_1 + \rho \cos \theta) + \bar{G}_{\mu\nu\rho k}^{(n)'} V_1) \\ &\quad (p^2 V_1^2 + 2p\rho V_1 \cos \theta + \rho^2)^{n-\mu} (pV_1 + \rho \cos \theta)^{\mu-\nu} V_1^{\mu+\nu}] . \end{aligned} \quad (4.3)$$

Considering third order terms only, and omitting primes and subscripts from the notation of (3.21)

$$\begin{aligned} \underline{\epsilon}_k &= \mu \{ [A(pV_1 + \rho \cos \theta) + \bar{A}V_1] [p^2 V_1^2 + 2p\rho \cos \theta + \rho^2] \\ &\quad + [B(pV_1 + \rho \cos \theta) + \bar{B}V_1] [pV_1 + \rho \cos \theta] \\ &\quad + [C(pV_1 + \rho \cos \theta) + \bar{C}V_1] V^2 \} + O(5). \end{aligned} \quad (4.4)$$

Using the identity $B = 2\bar{A}$, from M 20.42,

$$\begin{aligned} \underline{\epsilon}_k &= \mu \{ Ap^3 \cos \theta + [pA(1 + 2\cos \theta) + \bar{A}(2 + \cos 2\theta)] p^2 V_1 \\ &\quad + [3p^2 A + 6p\bar{A} + \bar{B} + C] p V_1^2 \cos \theta \\ &\quad + [p^3 A + 3p^2 \bar{A} + p(\bar{B} + C) + \bar{C}] V_1^3 \} + O(5). \end{aligned} \quad (4.5)$$

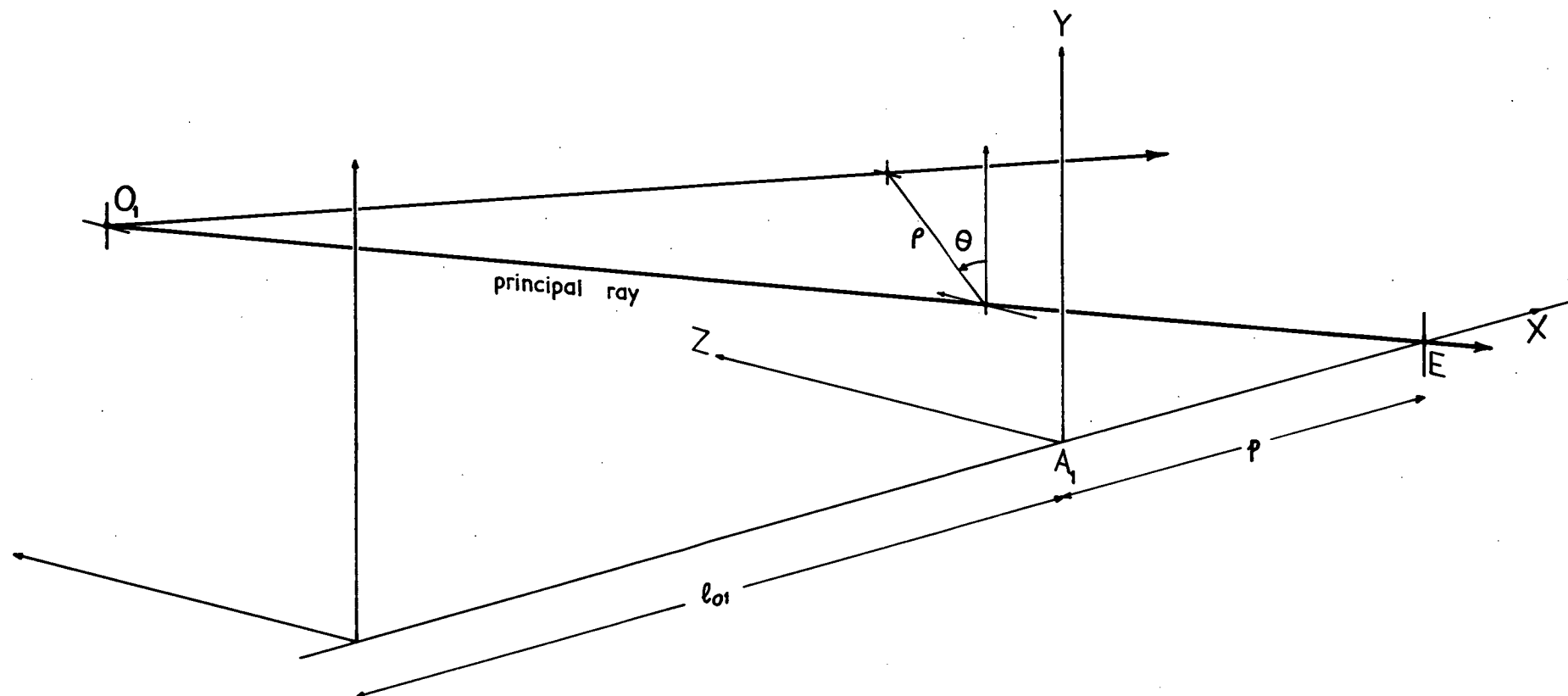


Fig. 5 Polar coordinates

If, now, $p = 0$

$$\epsilon_k = \mu \{ A \rho^3 \cos \theta + \bar{A} (2 + \cos 2\theta) \rho^2 v_1 + (\bar{B} + C) \rho v_1^2 \cos \theta + \bar{C} v_1^3 \} + O(5). \quad (4.6)$$

In this case, then, the p coefficients control the various well known types of aberrations in a simple manner. However, (4.6) results from the fact that the object is at infinity and the entrance pupil is at the first surface. In view of the complexity of (4.5) or (3.20), it is obviously of advantage if expansions similar to (4.6) could be used even when $v_{01} \neq 0$ and $p \neq 0$. By choosing a different set of initial coordinates to specify a ray this can be achieved. The new sets of coordinates are called "paracanonical coordinates". Paracanonical coordinates S_y, S_z, T_y, T_z , are defined in terms of the canonical coordinates by the following

$$\begin{aligned} \underline{S} &= \sigma \underline{Y}_1 + \bar{\sigma} \underline{V}_1 \\ \underline{T} &= \tau \underline{Y}_1 + \bar{\tau} \underline{V}_1, \end{aligned} \quad (4.7)$$

where $\sigma, \bar{\sigma}, \tau, \bar{\tau}$ are disposable constants subject only to the condition

$$g = \sigma \bar{\tau} - \bar{\sigma} \tau \neq 0. \quad (4.8)$$

Since $\underline{S}, \underline{T}$ are linearly dependent on $\underline{Y}_1, \underline{V}_1$, then any paraxial variable μ_j is related to the paraxial coordinates s, t by

$$\mu_j = \mu_{aj} s + \mu_{bj} t \quad (4.9)$$

$$\begin{aligned} \text{where } s &= \sigma y_1 + \bar{\sigma} v_1 \\ t &= \tau y_1 + \bar{\tau} v_1. \end{aligned} \quad (4.10)$$

The μ_{aj} , μ_{bj} are the paracanonical paraxial coefficients. The constants y_{aj} , y_{bj} , v_{aj} , v_{bj} are obtained in a similar manner to their canonical counterparts, except that here the paraxial tangential rays have as their initial data

$$\begin{aligned} \text{"a-ray"} \quad s &= 1, \quad t = 0 \\ \text{"b-ray"} \quad s &= 0, \quad t = 1. \end{aligned} \quad (4.11)$$

Thus, in terms of y_1 , v_1 , the starting data is, from (4.10)

$$\begin{aligned} \text{"a-ray"} \quad y_1 &= \bar{\tau}/g, \quad v_1 = -\tau/g \\ \text{"b-ray"} \quad y_1 &= -\bar{\sigma}/g, \quad v_1 = \sigma/g, \end{aligned} \quad (4.12)$$

whence, by using the recurrence relations (2.5) the values of y , v at each surface will be y_a , v_a for the a-ray and y_b , v_b for the b-ray.

The identity (2.10) becomes

$$y_a v_b - y_b v_a = N_1 / N g. \quad (4.13)$$

Thus an antepime associated with a symbol is now to be regarded as indicating division by N_1/g instead of N_1 as previously.

Since \underline{S} , \underline{T} are linearly dependent on \underline{Y}_1 , \underline{V}_1 , (3.14) can be written

$$\underline{\Delta\Delta} = \sum_{n=1}^{\infty} \sum_{\mu=0}^n \sum_{\nu=0}^{\mu} (g_{\mu\nu}^{(n)} \underline{S} + \bar{g}_{\mu\nu}^{(n)} \underline{T}) \xi_1^{n-\mu} \eta_1^{\mu-\nu} \zeta_1^{\nu}, \quad (4.14)$$

where

$$g_{\mu\nu}^{(n)} = g_{\mu\nu a}^{(n)} s + g_{\mu\nu b}^{(n)} t,$$

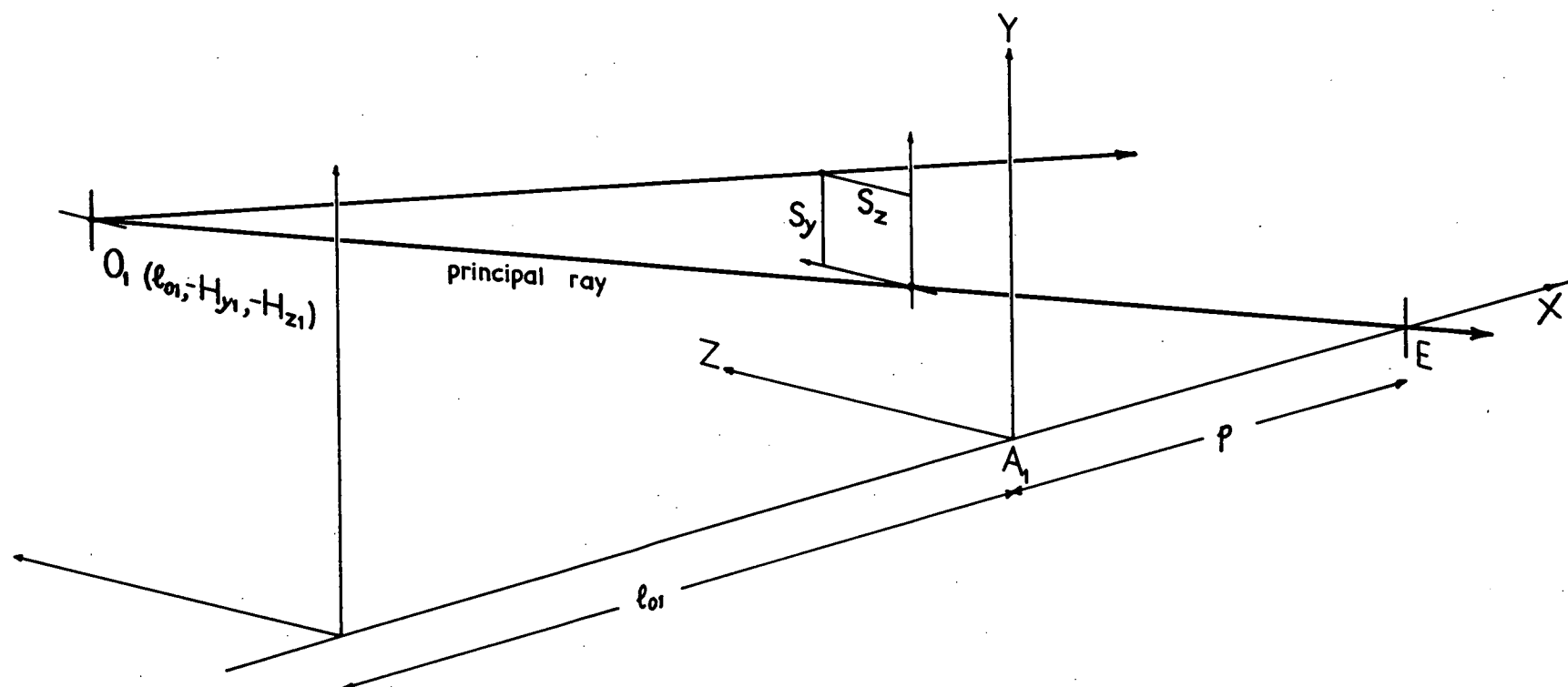


Fig. 6 Paracanonical $\bar{O}T$ coordinates
 $T_y = H_{y1}/l_{01}$, $T_z = H_{z1}/l_{01}$

and similarly for $\bar{g}_{\mu\nu}^{(n)}$.

Here

$$\begin{aligned}\xi_1 &= S_y^2 + S_z^2, \\ \eta_1 &= S_y T_y + S_z T_z, \\ \zeta_1 &= T_y^2 + T_z^2.\end{aligned}$$

(3.20) now becomes

$$\begin{aligned}\hat{e}^{\mu} &= \sum \sum \sum (G_{\mu\nu ak}^{(n)'} \underline{S} + \bar{G}_{\mu\nu ak}^{(n)'} \underline{T}) \xi_1^{\mu-\nu} \eta_1^{\mu-\nu} \zeta_1^{\nu} s_0 + \\ &+ \sum \sum \sum (G_{\mu\nu bk}^{(n)'} \underline{S} + \bar{G}_{\mu\nu bk}^{(n)'} \underline{T}) \xi_1^{\mu-\nu} \eta_1^{\mu-\nu} \zeta_1^{\nu} t_0.\end{aligned}\quad (4.15)$$

It is worthwhile noticing here that \underline{S} , \underline{T} , s_0 , t_0 have no surface subscripts since they are defined only at the first surface, see (4.7), that is, there are no variables S_j, T_j, s_{0j}, t_{0j} .

A particular set of paracanonical coordinates, $\bar{O}T$ coordinates, are defined as follows. Let the principal ray of a pencil intersect the first polar tangent plane in a point whose coordinates are \underline{Y}_e . Then S_y, S_z are defined as the y, z coordinates of the intersection of any ray of the pencil in this plane, taking \underline{Y}_e as origin. $T_y = H_{y1}/l_{01}$, $T_z = H_{z1}/l_{01}$ where l_{01} is the axial distance of the object plane from the first surface, see Fig.6. Thus, for these coordinates

$$\begin{aligned}\sigma &= l_{01}/(l_{01} - p), & \bar{\sigma} &= -pl_{01}/(l_{01} - p) \\ \tau &= -1/l_{01}, & \bar{\tau} &= 1, \\ g &= 1.\end{aligned}\quad (4.16)$$

From (4.10)

$$t_0 = \tau y_{01} + \bar{\tau} v_{01},$$

$$\begin{aligned}
 t_0 &= -y_{01}/l_{01} + v_{01} \\
 &\equiv 0.
 \end{aligned}
 \tag{4.17}$$

Thus the term involving t_0 in (4.15) is zero and does not appear in the description of the aberrations in this coordinate system, that is, only the "a" coefficients are required. Changing the polar coordinates,

$$\begin{aligned}
 S_y &= \rho \cos\theta, & S_z &= \rho \sin\theta \\
 T_y &= H_{y1}/l_{01} = \bar{H} \text{ say,} \\
 T_z &= 0,
 \end{aligned}$$

so that p does not appear explicitly. Then taking only 3rd order terms and omitting primes and subscript k from the coefficients, (4.15) becomes, taking ϵ'_{yk} for example

$$\begin{aligned}
 \epsilon'_{yk} &= \mu [A_a \rho^3 \cos\theta + \bar{A}_a (2 + \cos 2\theta) \rho^2 \bar{H} \\
 &\quad + (\bar{B}_a + C_a) \rho \bar{H}^2 \cos\theta + \bar{C}_a \bar{H}^3] + O(5),
 \end{aligned}
 \tag{4.18}$$

where $\mu = 1/Nk'v_{ak}$.

ϵ'_k is now described only by the a coefficients, and the form of the expansion is independent of the value of l_{01} or p , these being contained in the description of the particular coordinates. Note that this simplification comes about only if \underline{S} , \underline{T} are computed using the relevant values of l_{01} , p . A particular set of coefficients $A_a, \bar{A}_a \dots$ for one position of the object and entrance pupil cannot directly be used in (4.18) to describe the aberrations for a different object or entrance pupil plane. It is possible to calculate, from a particular set of paracanonical aberration coefficients, other sets of coefficients corresponding to different situations, see M Section 32, but this may not be simple.

This simplification of the form of expressions which arises from the use of paracanonical coordinates is further demonstrated in the following example.

From (3.7)

$$\begin{aligned} \underline{h}_k &= m_k' \underline{h}_1 \\ &= N_1 v_{01} \underline{h}_1 / N_k' v_{0k}' \\ &= N_1 \underline{V}_1 (\underline{l}_{01} - p) / N_k' (v_{0k}' \underline{l}_{01} + v_{0k}') , \end{aligned} \quad (4.19)$$

where \underline{V}_1 refers to the principal ray.

The relation between the canonical and paracanonical coefficients is given by

$$\begin{aligned} \mu_{pj} &= \sigma \mu_{aj} + \tau \mu_{bj} , \\ \mu_{qj} &= \bar{\sigma} \mu_{aj} + \bar{\tau} \mu_{bj} . \end{aligned} \quad (4.20)$$

With these and using $\bar{O}T$ coordinates it is easy to show that (4.19) becomes

$$\underline{h}_k = N_1 \underline{T} / N_k' v_{0k}' . \quad (4.21)$$

It has already been mentioned that $t_0 \equiv 0$ for $\bar{O}T$ coordinates (4.17), which leads to considerable simplification in the computations required to obtain the aberration coefficients. This feature is common to all the sets of paracanonical coordinates described in detail in M Section 13 (a), (b), (c) and (d). These coordinates also possess an additional feature, namely that $\underline{T} (= -\tau \underline{H}_1, \text{ M 13.71})$ is a constant for all rays that issue from a particular object point. Thus in any pencil, the only coordinates which vary are S_y, S_z , in contrast with canonical coordinates, in which, for a near point object, Y_1, Z_1, V_1, W_1 all vary from ray to ray. This feature is of considerable

advantage when spot diagrams are to be computed.

$\bar{O}T$ coordinates, in particular, have further advantages over the other paracanonical coordinates. Firstly, they will apply when the object is at infinity, in which case $\underline{T} = \underline{V}_1$. Secondly, $g = 1$, and thus the factor N_1/g is usually unity since most optical systems considered work in air. This means that the antéprimes which occur in the iteration formulae (see Section 5) may be omitted during computation of the coefficients. If N_1 is not unity, say in an oil-immersion lens, then it can be made unity by multiplying all refractive indices by $1/N_1$, the image forming properties of the optical system being unchanged by this process.

In view of (2.6), (4.10); (3.14), (4.14); and (2.14), (4.13), it is obvious that the derivation of the formulae for calculating aberration coefficients using either canonical or paracanonical coordinates will be formally identical. Consequently, to avoid a multiplicity of symbols, canonical symbols will be used almost entirely in the following work. If paracanonical coordinates are being considered, then, of course

$$\begin{aligned}\underline{Y}_1 & \text{ stands for } \underline{S} , \\ \underline{V}_1 & \text{ stands for } \underline{T} , \\ \xi_1 & \text{ stands for } (S_y^2 + S_z^2) , \\ \eta_1 & \text{ stands for } (S_y T_y + S_z T_z) , \\ \zeta_1 & \text{ stands for } (T_y^2 + T_z^2) ,\end{aligned}\tag{4.22}$$

the subscripts p, q are to be replaced by a, b , and the antéprimes will mean division by N_1/g . It is worthwhile here to repeat that $\underline{S}, \underline{T}$ are coordinates defined at the first surface only that is, there are no variables S_j, T_j .

Paracanonical symbols will, in general, be used to indicate that only these coordinates are being considered. e.g. (4.21).

5. Calculation of Aberration Coefficients

(a) Iteration

The formulae for calculating the aberration coefficients are obtained by iterative processes. To demonstrate this process two examples will be used. The first of these indicates the basic method of the process. The second example is a simplification of the process as carried out in M, so that by an extension to a number of quantities, equations of the type M 11.3 can be obtained. This example also introduces the idea of "pseudo-expansion" and "intrinsic coefficients".

Example 1

Suppose that x is some function $f(t)$ of t , and that x and t are related by the equation

$$x^2 - x + t = 0. \quad (5.1)$$

We shall assume that $f(t)$ is a power series in ascending powers of t , and is valid for all values of x . Suppose also that as $x \rightarrow 0$, $t \rightarrow 0$. From this we can deduce that there are no constant terms in $f(t)$.

Rewriting (5.1), we get

$$x = t + x^2. \quad (5.2)$$

Then, as x becomes small, to a first approximation

$$x = t. \quad (5.3)$$

Substituting (5.3) in (5.2), one obtains

$$x = t + t^2, \quad (5.4)$$

thus we have a better approximation of $f(t)$.

Reinserting (5.4) in (5.2),

$$\begin{aligned} x &= t + (t + t^2)^2 \\ &= t + t^2 + 2t^3 + O(4). \end{aligned} \quad (5.5)$$

By continuing this process of substitution, the series for $f(t)$ can be determined. This process, by which more accurate approximations to the series for x are obtained by successive back substitution is called iteration. Iteration will only work when, as a result of resubstitution, the new terms generated are of higher order than those that existed before substitution.

For example, suppose that $x \rightarrow 1$ as $t \rightarrow 0$, in conjunction with (5.1). Then to a first approximation

$$x = 1 + t. \quad (5.6)$$

Substituting (5.6) in (5.2), we get

$$\begin{aligned} x &= t + (1 + t)^2 \\ &= 1 + 3t + t^2. \end{aligned} \quad (5.7)$$

Resubstitution of (5.7) in (5.2) gives

$$x = 1 + 7t + 9t^2 + 6t^3 + O(4). \quad (5.8)$$

It will be noticed that the coefficients of $t^0, t^1, t^2 \dots$ will now vary at each iterative step so that the process does not converge to a fixed series for x . This is a result of the constant term in (5.6), (5.7), (5.8).

Consider once again the series (5.5). This is developing as

$$x = t + t^2 + 2t^3 + 5t^4 + o(5). \quad (5.9)$$

Suppose we write the series for x as

$$x = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + \dots \quad (5.10)$$

Now, since $x \rightarrow 0$ as $t \rightarrow 0$,

$$a_0 = 0, \quad (5.11)$$

so that (5.10) can now be written

$$x = a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + o(5). \quad (5.12)$$

Substituting (5.12) in (5.2), (5.2) becomes

$$\begin{aligned} x &= t + (a_1 t + a_2 t^2 + a_3 t^3 + \dots) \\ &= t + a_1^2 t^2 + 2a_1 a_2 t^3 + (2a_1 a_3 + a_2^2) t^4 + o(5). \end{aligned} \quad (5.13)$$

Comparing coefficients between (5.12) and (5.13), one finds that

$$\begin{aligned} a_1 &= 1 \\ a_2 &= a_1^2 = 1 \\ a_3 &= 2a_1 a_2 = 2 \\ a_4 &= (2a_1 a_3 + a_2^2) = 5 \end{aligned} \quad (5.14)$$

thus (5.12) can be written

$$x = t + t^2 + 2t^3 + 5t^4 + o(5), \quad (5.15)$$

which is identical with (5.9). Thus the "complete" series for x can be obtained in one operation. This method will, of course,

only give correct results in those cases where direct iteration produces a unique series for x (in contrast with (5.7), (5.8)).

It will be noticed in (5.14) that the coefficient a_n is expressed as some combination of coefficients of terms of degree lower than n . Thus having found the first coefficient explicitly, the rest can be easily computed.

Example 2

Consider the series for $\Delta\Delta$, (3.14), but for the sake of simplicity suppose that there is only one variable Y , so that now

$$\Delta\Delta = \sum_{n=1}^{\infty} g^{(n)} Y_1 \xi_1^n, \quad (5.16)$$

where $\xi_1 = Y_1^2$.

As before, we shall write the first three terms of (5.16) explicitly as

$$\Delta\Delta = aY_1\xi_1 + sY_1\xi_1^2 + tY_1\xi_1^3 + O(9), \quad (5.17)$$

that is,

$$g^{(1)} = a, \quad g^{(2)} = s, \quad g^{(3)} = t.$$

Now, at any surface

$$\Delta\Delta_j = f(Y_j), \quad \text{cf}(3.13). \quad (5.18)$$

Suppose that

$$Y_j = \bar{k}(Y_1 + \sum_1^{j-1} \Delta\Delta_1), \quad (5.19)$$

where \bar{k} is a constant of the surface j , depending only on the

constitution of the optical system.

Let the series for $\sum_1^{j-1} \Delta\Lambda_1$ be written

$$\sum_1^{j-1} \Delta\Lambda_1 = A_j Y_1 \xi_1 + S_j Y_1 \xi_1^2 + T_j Y_1 \xi_1^3 + O(9), \quad (5.20)$$

where

$$A_j = \sum_1^{j-1} a_1,$$

$$S_j = \sum_1^{j-1} s_1,$$

$$T_j = \sum_1^{j-1} t_1, \dots\dots$$

Since Λ is an optical invariant when Y becomes small, to a first approximation

$$Y_j = \bar{k} Y_1. \quad (5.21)$$

Inserting (5.21) in (5.18), $\Delta\Lambda_j = f(\bar{k} Y_1)$ is a first approximation to $\Delta\Lambda_j$. Knowing the function $f(\bar{k} Y_1)$, a series for $\Delta\Lambda_j$ in terms of Y_1 can be formed similar to (5.17), namely

$$\begin{aligned} \Delta\Lambda_j &= \sum_1^{\infty} g_1^{(n)} Y_1 \xi_1^n \\ &= \underline{a}_j Y_1 \xi_1 + \underline{s}_j Y_1 \xi_1^2 + \underline{t}_j Y_1 \xi_1^3 + O(9) \end{aligned} \quad (5.22)$$

This series is not the one required since the term $\sum_1^{j-1} \Delta\Lambda_1$ has been omitted from the expression for Y (5.19). The series (5.22) is called the "pseudo-expansion" of $\Delta\Lambda_j$. The coefficients of this series $\underline{a}_j, \underline{s}_j, \underline{t}_j \dots$ depend only on \bar{k} , and hence they also depend only on the construction of the system. For this

reason they are called "intrinsic coefficients".

Now (5.22) gives us $\Delta\Lambda_j$ correct to the first order, cf (5.3), hence (5.20) is also known correct to the first order. Thus the substitution of $(Y_i + \sum_{j=1}^{j-1} \Delta\Lambda_j)$ for Y_i in (5.22) will give $\Delta\Lambda_j$ correct to the second order, since $\sum \Delta\Lambda$ is of $O(3)$ and will not affect the terms already present in (5.22), cf (5.4). Thus we have $\sum \Delta\Lambda$ correct to the second order, and the process of resubstitution in (5.22) can be continued, and so the series for $\Delta\Lambda_j$ can be obtained.

As before, the process can be condensed by writing the correct series for $\Delta\Lambda_j$ as

$$\Delta\Lambda_j = a_j Y_i \xi_i + s_j Y_i \xi_i^2 + t_j Y_i \xi_i^3 + O(9) \quad (5.23)$$

and also

$$\sum_{j=1}^{j-1} \Delta\Lambda_j = A_j Y_i \xi_i + S_j Y_i \xi_i^2 + T_j Y_i \xi_i^3 + O(9), \quad (5.24)$$

then, by replacing Y_i in (5.22) by $(Y_i + \sum_{j=1}^{j-1} \Delta\Lambda_j)$ and comparing coefficients between the resultant series and (5.23), the formulae for obtaining the higher order coefficients can be read off directly. (By way of comparison, in example 1, $x = t$, (5.3), corresponds to the pseudo-expansion of $\Delta\Lambda_j$, t corresponds to $f(\bar{k}Y_i)$, x to $\Delta\Lambda_j$ and x^2 in (5.2) corresponds to $f(\bar{k} \sum_{j=1}^{j-1} \Delta\Lambda_j)$.)

Thus

$$\begin{aligned}
\Delta\Lambda &= \underline{a}(Y_1 + AY_1\xi_1 + SY_1\xi_1^2 + TY_1\xi_1^3 + \dots)(\xi_1 + 2Y_1[AY_1\xi_1 + SY_1\xi_1^2 + \dots] + \\
&\quad + [AY_1\xi_1 + SY_1\xi_1^2 + \dots]^2) \\
&\quad + \underline{s}(Y_1 + AY_1\xi_1 + SY_1\xi_1^2 + \dots)(\xi_1 + 2Y_1[AY_1\xi_1 + \dots] + [AY_1\xi_1 + \dots]^2)^2 \\
&\quad + \underline{t}(Y_1 + \dots)(\xi_1 + \dots)^3 + O(9) \\
&= a(Y_1 + AY_1\xi_1 + SY_1\xi_1^2 + \dots)(\xi_1 + 2A\xi_1^2 + 2S\xi_1^3 + \dots + A^2\xi_1^3 + \dots) \\
&\quad + s(Y_1 + AY_1\xi_1 + \dots)(\xi_1^2 + 4A\xi_1^3 + \dots) \\
&\quad + t(Y_1 + \dots)(\xi_1^3 + \dots) + O(9) . \tag{5.25}
\end{aligned}$$

Comparing coefficients between (5.23) and (5.25), we get from the coefficient of $Y_1\xi_1$

$$a = \underline{a} , \tag{5.26}$$

$Y_1\xi_1^2$

$$\begin{aligned}
s &= s + (2A + A)\underline{a} \\
&= \underline{s} + 3aA, \tag{5.27}
\end{aligned}$$

$Y_1\xi_1^3$

$$\begin{aligned}
t &= \underline{t} + \underline{s}(4A + A) \\
&\quad + \underline{a}(2S + A^2 + 2A^2 + S) \\
&= \underline{t} + 5sA + 3a(S + A^2) . \tag{5.28}
\end{aligned}$$

(5.26), (5.27), (5.28) are the iteration formulae which enable the contributions a, s, t to the aberration coefficients, (5.23), at any surface to be obtained from the intrinsic coefficients of the surface and the sum of the contributions over the previous surfaces, (5.24).

For example, at the first surface

$$A_1 = S_1 = T_1 = 0,$$

since there are no previous surfaces over which to sum contributions,

hence, from (5.26), (5.27), (5.28) one finds that

$$\begin{aligned} a_1 &= \underline{a}_1 \\ s_1 &= \underline{s}_1 \\ t_1 &= \underline{t}_1 . \end{aligned}$$

Now

$$\begin{aligned} A_2 &= a_1 \\ S_2 &= s_1 \\ T_2 &= t_1 , \end{aligned}$$

so that at the second surface

$$\begin{aligned} a_2 &= \underline{a}_2 \\ s_2 &= \underline{s}_2 + 3\underline{a}_2 a_1 \\ t_2 &= \underline{t}_2 + 5\underline{s}_2 a_1 + 3\underline{a}_2 (s_1 + a_1^2) . \end{aligned}$$

Then

$$\begin{aligned} A_3 &= a_1 + a_2 \\ S_3 &= s_1 + s_2 , \\ &\vdots \end{aligned}$$

and the coefficients a_3, s_3, t_3 can be found.

In this manner the contributions to the final aberration coefficients by each surface are found. Notice here that any inaccuracy in the value of ΔA_j and hence of \underline{a}_j comes, not from the coefficients, which are exact, but from the fact that we terminate the infinite series for $\Delta A_j(\underline{a}_j)$ without regard to the remaining terms.

In M, iteration is carried out with four variables at each surface, namely Y_j, Z_j, V_j, W_j , and corresponding to these, four coordinates at the first surface, Y_1, Z_1, V_1, W_1 . The

relations between these, equivalent to (5.19) are, from M 9.3.4,

$$\underline{Y}_j = y_{pj} (\underline{Y}_1 + \underline{\delta}_{y1}) + y_{qj} (\underline{V}_1 + \underline{\delta}_{v1}) \quad (5.29)$$

$$\underline{V}_j = v_{pj} (\underline{Y}_1 + \underline{\delta}_{y1}) + v_{qj} (\underline{V}_1 + \underline{\delta}_{v1}),$$

where

$$\begin{aligned} \underline{\delta}_{y1} &= - \sum_{i=1}^{j-1} \Delta \underline{\Lambda}_{qi}, \\ \underline{\delta}_{v1} &= + \sum_{i=1}^{j-1} \Delta \underline{\Lambda}_{pi} \quad (\text{note antepimes}). \end{aligned} \quad (5.30)$$

Each of the relations (5.30) represents two relations, e.g.

$$\begin{aligned} \delta_{y1} &= - \sum_{i=1}^{j-1} \Delta \underline{\Lambda}_{yqi}, \\ \delta_{z1} &= - \sum_{i=1}^{j-1} \Delta \underline{\Lambda}_{zqi}. \end{aligned} \quad (5.31)$$

$\Delta \underline{\Lambda}_q$ is, of course, given by an expression of the form (3.14), but containing only q coefficients, (3.19), i.e. $y_{0i} = 0$, and similarly for $\Delta \underline{\Lambda}_p$.

If now "increments" $\underline{\delta}_y$, $\underline{\delta}_v$ are neglected in (5.29) i.e.

$$\underline{Y}_j = y_{pj} \underline{Y}_1 + y_{qj} \underline{V}_1, \quad (5.32)$$

$$\underline{V}_j = v_{pj} \underline{Y}_1 + v_{qj} \underline{V}_1$$

then using (5.32) in (3.13) or an equivalent expression, the pseudo-expansion of $\Delta \underline{\Lambda}$ will be obtained, see M 10.1. If then in the pseudo-expansion \underline{Y}_1 , \underline{V}_1 are replaced by $(\underline{Y}_1 + \underline{\delta}_y)$, $(\underline{V}_1 + \underline{\delta}_v)$ the iteration formulae M 10.2, 11.3, 81.3 can be obtained. The intermediate p and q coefficients in M 11.3, 81.3 arise of course from $\underline{\delta}_v$ and $\underline{\delta}_y$.

6. Intrinsic Coefficients

The straightforward method of calculating the intrinsic coefficients is to expand the expression (3.13) for $\Delta\Lambda$ as a power series in $\underline{Y}_j, \underline{V}_j, \xi_j, \eta_j, \zeta_j$, then convert this series to one in $\underline{Y}_1, \underline{V}_1, \xi_1, \eta_1, \zeta_1$ by means of (5.32), whence the coefficients of the resultant series are the intrinsic coefficients. However, it proves more convenient to first factorize (3.13) and then expand one of the factors as a power series. Multiplying the coefficients of this series by the appropriate factor then gives the required coefficients.

Now

$$\Delta\Lambda = \underline{E} J, \quad (6.1)$$

from M 14.5, where $\underline{E} = N r \underline{I}$ and J is given by M 15.1

$$\text{namely } J = u_0(\underline{S} - 1) - v_0(\underline{T} - 1). \quad (6.2)$$

Taking first order terms only,

$$\begin{aligned} \Delta\Lambda_{\text{int}}^{(1)} &= \underline{E}_{\text{int}} J_{\text{int}}^{(1)} \\ &= (e_p \underline{Y}_1 + e_q \underline{V}_1) J_{\text{int}}^{(1)}, \end{aligned} \quad (6.3)$$

where $e_p = N r_{1p}$, $e_q = N r_{1q}$. Here \underline{E} has already been expressed in terms of $\underline{Y}_1, \underline{V}_1$, by means of (5.32) (i.e. (5.29) neglecting increments).

Now, for example, a is the coefficient of $\underline{Y}_1 \xi_1$, that is the product, $e_p \cdot (\text{coeff. of } \xi_1 \text{ in } J_{\text{int}}^{(1)})$. (6.4)
Without going through the algebra, M 17.4 gives the expression for $J^{(1)}$,

$$J^{(1)} = u_0 \gamma_1 + v_0 \gamma_2. \quad (6.5)$$

where the γ 's are given by M 17.1, namely

$$\begin{aligned}\gamma_1 &= \frac{1}{2}(k^2 - k + 1)c^2\xi - k^2 c\eta + \frac{1}{2}k(k+1)\zeta \\ \gamma_2 &= -\frac{1}{2}(k-1)^2 c^2\xi + k(k-1)c\eta + \frac{1}{2}(1-k^2)\zeta,\end{aligned}\quad (6.6)$$

where $k = N/N'$. Since $u_a = \Delta v_a = v'_d - v_a$, (6.5) becomes

$$J^{(1)} = (v'_d - v_a)\gamma_1 + v'_d\gamma_2. \quad (6.7)$$

Now, in (6.6)

$$\begin{aligned}\xi &= Y_j^2 + Z_j^2 \\ \eta &= Y_j V_j + Z_j W_j \\ \zeta &= V_j^2 + W_j^2,\end{aligned}$$

therefore, expressing ξ, η, ζ , in terms of $\underline{Y}, \underline{V}$, by means of (5.32), we get

$$\begin{aligned}\xi &= y_p^2 \xi_1 + 2y_p y_q \eta_1 + y_q^2 \zeta_1 \\ \eta &= y_p v_p \xi_1 + (y_p v_q + y_q v_p) \eta_1 + y_q v_q \zeta_1 \\ \zeta &= v_p^2 \xi_1 + 2v_p v_q \eta_1 + v_q^2 \zeta_1.\end{aligned}\quad (6.8)$$

Thus the coefficient of ξ_1 in γ_1 is

$$\frac{1}{2}[(k^2 - k + 1)c^2 y_p^2 - 2k^2 c y_p v_p + k(k+1)v_p^2]. \quad (6.9)$$

Now, it can easily be shown that

$$v_p'^2 = (k^2 - 2k + 1)c^2 y_p^2 + 2k(1-k)c y_p v_p + k^2 v_p^2 \quad (6.10)$$

and

$$i_p i_p' = k(c^2 y_p^2 - 2c y_p v_p + v_p^2) . \quad (6.11)$$

From (6.10), (6.11) it will be seen that (6.9) can be written

$$\frac{1}{2}(i_p i_p' + v_p'^2) . \quad (6.12)$$

Similarly, the coefficient of ξ_1 in γ_2 is

$$\begin{aligned} & \frac{1}{2}[-(k-1)^2 c^2 y_p^2 + 2k(k-1)c y_p v_p + (1-k^2)v_p^2] \\ &= \frac{1}{2}[v_p^2 - \{(k^2 - 2k + 1)c^2 y_p^2 + 2k(1-k)c y_p v_p + k^2 v_p^2\}] \\ &= \frac{1}{2}(v_p^2 - v_p'^2) . \end{aligned} \quad (6.13)$$

Thus, from (6.4), (6.7), (6.12), (6.13)

$$\begin{aligned} \underline{a} &= \frac{1}{2} N r i_p [(v_d - v_o)(i_p i_p' + v_p'^2) + v_o(v_p^2 - v_p'^2)] \\ &= \frac{1}{2} N r i_p [\{(v_p' - v_p) y_{o1} + (v_d - v_o) v_{o1}\} \{i_p i_p' + v_p'^2\} + \{v_p y_{o1} + v_o v_{o1}\} \\ &\quad \{v_p^2 - v_p'^2\}] . \end{aligned}$$

Now, \underline{a}_p is the coefficient of $y_{o1} Y_1 \xi_1$ (\underline{a}_q is the coefficient of $v_{o1} Y_1 \xi_1$), thus

$$\begin{aligned} \underline{a}_p &= \frac{1}{2} N r i_p [(v_p' - v_p)(i_p i_p' + v_p'^2) + v_p(v_p^2 - v_p'^2)] \\ &= \frac{1}{2} N r i_p (i_p - i_p') [(i_p i_p' + v_p'^2) - v_p'(v_p + v_p')] \\ &= \frac{1}{2} N (1 - k) y_p i_p^2 (i_p' - v_p') , \quad \text{see M 24.1} . \end{aligned} \quad (6.14)$$

Since the barred coefficients are associated with \underline{V}_1 , then \bar{a}_p, \bar{a}_q are, from (6.3), eq. (coeff. of y_0, ξ_1, v_0, ξ_1 in $J^{(1)}$). In this manner, the coefficients in M 24.2, 24.3, 218.7 are obtained.

PART II - Computer and Programmes7. Computer Description

The computer used for this work is the English Electric DEUCE housed at the University of New South Wales, Sydney. The machine operates entirely in the serial mode on 32 bit numbers. The high speed store consists of acoustic mercury delay lines of various lengths, the smallest containing one word (32 bits) and the largest containing 32 words. In detail, the high speed storage is as follows:

- 1) Four single word stores.
- 2) Three double word stores.
- 3) Two four word stores.
- 4) Twelve 32 word stores.

the total high speed storage being 402 words. The slow speed backing store is a magnetic drum containing 8192 words arranged in 256 tracks of 32 words each. The drum has separate read and write heads on opposite sides of the drum, each bank consisting of 16 heads which can be located vertically in any one of 16 positions, thus covering the entire 256 tracks. Information is read to or from the drum via one of the 32 word delay lines in blocks of 32 words. Drum transfers and head shifts proceed automatically once they are set up and do not interfere with the normal operation of the machine, provided that access to the transfer delay line is not required. The transfer of a block of words from the drum takes about 13 m secs, and a head shift about 50 m secs. The digit frequency of the machine is 1 megacycle; thus a single word store containing 32 bits needs 32 μ secs to deliver one complete word at its output. Hence the basic timing of the machine is tied to this time which is called a "minor cycle" abbreviated m.c. The circulation time of a 32 word store is 32 m.c., that is, the same word is presented at the output every 1024 μ secs (approximately 1 msec). This interval of time is termed a "major cycle", abbreviated M.c. The 32 word stores are termed "delay lines". Thus the machine contains 12 delay lines, numbered for addressing purposes from 1 to 12. The remaining shorter delay lines are simply designated as single, double, or quadruple stores.

It is a feature of this machine that arithmetic processes are associated with four stores, three of these being single stores, and one a double store. The way in which these stores are

addressed determines the type of arithmetic operation performed. The inputs and outputs of all high speed stores are connected via numbered gates to a common line called the "main highway". The output gates are called "source gates" or more simply "sources", and the input gates are called "destinations". Thus the basic instruction of the machine contains two addresses, one for the source and one for the destination. It will be realized that if the input and output of any one store are connected, the word contained in the store will circulate indefinitely. A non-destructive read out is obtained simply by tapping this circulation path. The read in of a new word is obtained by breaking the circulation path and connecting the input to the main highway, the circulation path being restored as soon as the last digit is read in to retain the word in the store (the connection to the main highway being opened at the same time of course).

Owing to the fact that words in the delay lines appear sequentially at the source gate, the instruction word contains an extra number, called the wait number, which determines at what m.c. the instruction will be obeyed. Since the circulation time of the longest store is 32 m.c., numbers in the range 0-31 are sufficient for this purpose. Zero time is usually taken to be the m.c. in which the first instruction/word enters a delay line from the card reader at initial input. Thus words in a delay line are denoted by 0, 1, ... 31, "0" referring to the first word read into the machine in any particular programme. Programme instructions/words are read into the delay lines from cards in binary form, and the input programme is so arranged that, having read the first word into the first delay line, the first word read into any other delay line also has a wait number of 0 associated with it. Thus the subsequent location of words in any delay line is known.

The control section of the machine which interprets instructions and sets up the required gates has direct access to any of the first eight delay lines. However, any particular instruction in a delay line is not accessible to control except at the output. Thus an additional number is required in the instruction word which indicates to the control when the required next instruction is accessible, called a timing number. As with the wait number, the timing number is in the range 0-31. The particular delay line is also specified, of course, and the numbers here are in the range 0-7, 0 referring to delay line 8. Finally, in the instruction word, provision is made for continuing a transfer between stores for more than 1 m.c. It is possible, in fact, to

transfer the 32 words of one delay line to another with one instruction. With one instruction transfers can be performed for from 1-32 m.c. The duration of a transfer is determined by the difference between the wait and timing numbers, which, in the case of long transfers (more than one word) restricts the location of the next instruction. The instruction word is as follows:

Digit position in word	Function
1 (least significant digit)	Not used
2	Next instruction source, i.e. Delay line 0-7
3	
4	
5	Source gate number Range 0-31
6	
7	
8	
9	
10	Destination gate number Range 0-31
11	
12	
13	
14	
15	Long or short transfers
16	
17	Wait number, referred to time instruction enters control
18	
19	
20	
21	
22	Special use
23	
24	
25	
26	Timing number, referred to time instruction enter control
27	
28	
29	
30	
31	Not used
32	Stop or go

It will be noticed in the preceeding table that the wait and timing numbers refer to the time that the instruction containing them enters control as their time reference. As mentioned before, the machine keeps track of m.c. only, hence it does not know which m.c. of a delay line the programmer has referred to as 0. Thus it is the programmer's job to keep track of the location of words at any subsequent time. For instance, if a word with the m.c. number 6 enters control, and a transfer of word number 27 is required, then the wait number in the instruction will be 19 since a minimum 2 m.c. are required to obey an instruction - one m.c. to read in the instruction and one m.c. to set up the gates. Then, after waiting 19 m.c. from the set up time, the gates are opened and word 27 enters the main highway. The minimum time to obey any instruction is 2 m.c. e.g. an instruction entering control in m.c. 11 can, at best, be obeyed in m.c. 13, the wait number for this being 0. Thus the sequence of events here are, instruction enters control in m.c. 11, gates set up in m.c. 12, no wait so obeyed in m.c. 13. The location of words by the programmer is facilitated by arranging the coding sheets in columns of 32 words representing the state of a delay line at m.c. times 0, 32, 64, 96, ... Apart from the direct access to delay lines 1-8, control can accept instructions from any other high speed storage location, but this requires a separate instruction, this being considered as simply a normal transfer between stores. In this event, the next-instruction-source number of the instruction is ignored by control, since, when the transfer is complete, control has its next instruction.

The DEUCE machine has no indexing registers, but digits 22-25 in the instruction word may be used for this purpose. Digits in these locations are always ignored by control. However, by placing an instruction in one of the stores that has adding facilities, digits may be added into these locations each time the instruction is obeyed during a repetitive loop. Eventually, these will overflow into position 26, that is, the timing number will be increased by 1 and a different instruction from the normal one of the loop will enter control and thus exit can be made from the loop. A particular use of this facility is when numbers are being transferred successively from a delay line by increasing the wait number each time around the loop. If the digits 22-25 are all ones, when the wait number exceeds 31, the overflow will be carried right through to the timing number, giving exit to the loop.

As mentioned before, the minimum time to obey an instruction is 2 m.c. (64 μ sec), but, owing to the facility of being

able to transfer words for more than 1 m.c. with one instruction, it is possible to add/subtract the entire contents of a delay line (32 words) in 33 m.c. For example, words sent to destination 25 are added to the contents of the single word store 13 (see Fig. 7). Thus, if 13 is originally clear, then a long transfer of 32 words from delay line 1, say, to destination 25 will produce in 13 the sum of the contents of delay line 1 after 33 m.c. from the time the instruction entered control. It will be seen in Fig. 7, the schematic diagram for DEUCE that double word store 21 also has adding/subtracting facilities. However, this store is not generally used for sequential adding, these facilities are to allow arithmetic operations on the results of multiplication/division, which appear in this store. Multiplication/division are carried out on numbers stored in 16, 21, and are initiated by "trigger" instructions. Shift facilities are associated with store 14, the digits appearing at source 23 are those of 14 shifted down one place, the digits from 24 are those of 14 shifted up one place. Multiple shifts can be simply obtained by a single instruction of the type 23-14 for any number of m.c. up to 32, whence the digits of 14 will be shifted as many places as the number of m.c. of the transfer.

Logical operations are performed between the contents of stores 14, 15 by sources 25, 26. Discrimination on sign or magnitude of a number are performed by destinations 27, 28 respectively (negative numbers are two's complement). If a positive number is sent to destination 27, then the next instruction as indicated by the timing number is taken into control; if the number sent to 27 is negative, then the timing number is increased by 1, causing control to take a different instruction.

It is possible to alter the m.c. numbers of words in a long delay line by using delay line 10 and store 16. By initiating a special trigger (T.C.A.) the output of 10 is connected to the input of 16. Then a long transfer for 32 m.c. between source 16 and destination 10 will result in the words in 10 being shifted down one m.c., i.e. word number 0 becomes word number 1

word number 31 becomes word number 0.

Sources 27-31 are constants, 28 generates wait numbers and is used mainly for instruction modification. Source 30 is used to clear stores. Source 31 is mainly used in logical operations.

In normal operations, instructions are punched in binary on cards, 12 instructions per card, 3 cards per delay line. Thus at a rate of 200 cards per minute instructions are read in at the

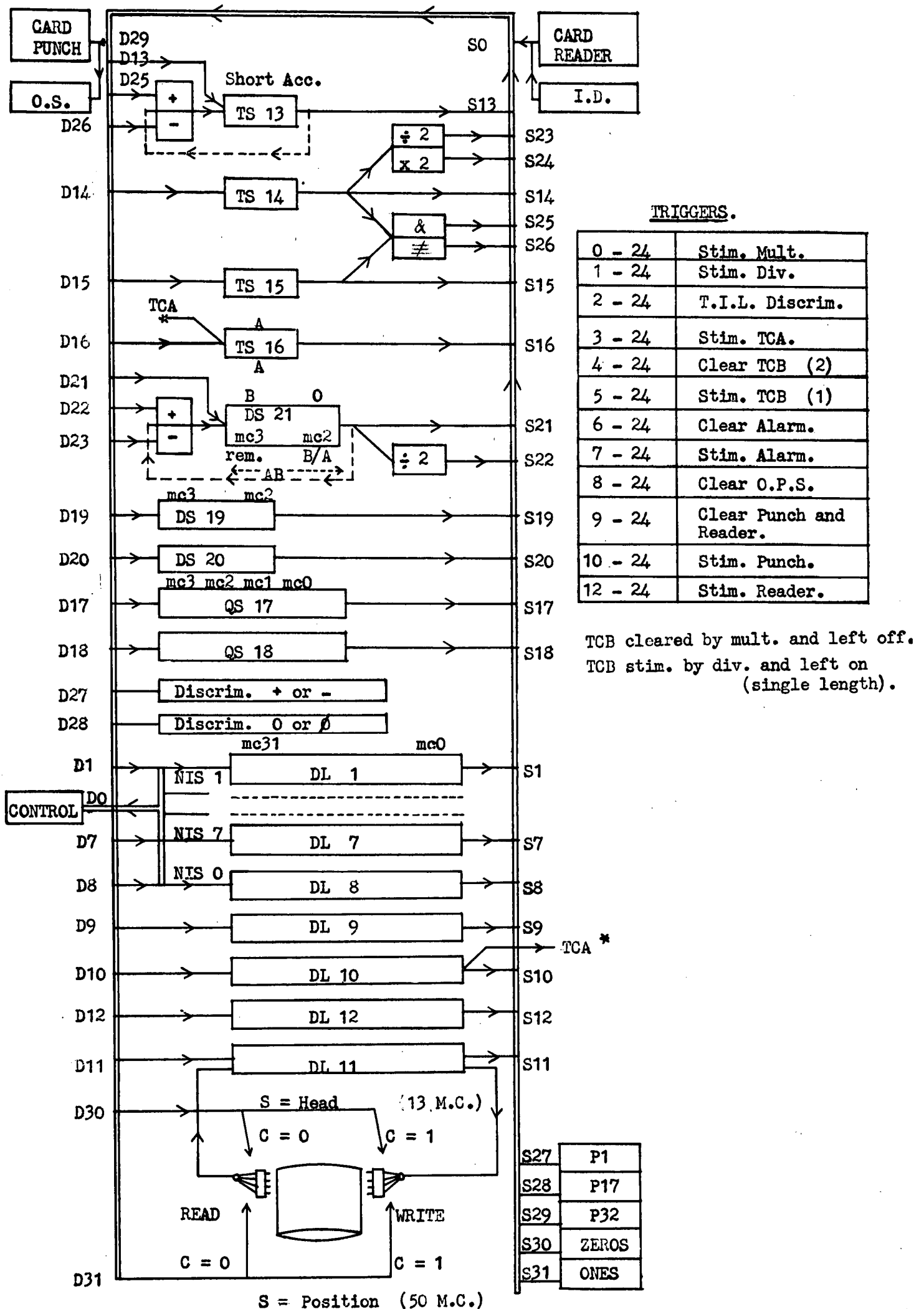


Fig. 7 DEUCE Block diagram

rate of 2,400 per minute. In most cases, numerical data are decimally punched and converted to binary by subroutine while being read at the normal card-reading rate. The cards used are 80 column cards of which 32 columns are read by the machine. Punching out is at the rate of 100 cards per minute, decimal punching being achieved by programme.

DEUCE has a built-in programme testing facility of considerable power. When a programme is initially written, "stops" are inserted at intervals by simply blocking up hole 32 in the appropriate instructions on the cards. When the machine encounters an instruction without a "go" digit it will halt, pending an external stimulus. If it is known that an error is present between any pair of "stoppers", it is possible to trigger the machine to punch out the instructions as they are obeyed, whence they can be compared with the detailed flow and coding sheets. It will be realized that instructions in DEUCE are not stored in the order in which they are obeyed, they are fitted into the nearest vacant word location so as to minimise access time. Thus a "store dump" does not indicate the order in which instructions will be obeyed. The facility of punching out instructions as they are obeyed is particularly useful when discrimination is involved, as this can be forced by manual control for testing purposes. The causes of closed loops can also be detected with ease by this facility, especially in those cases where the loop is unintentional, i.e. caused by instruction overwriting.

8. Aberration Coefficients Programme

The first programme produced was for the computation of the aberration coefficients and the contributions to them by the individual surfaces of a system containing 16 surfaces or less. The programme is based on a published table⁴ for the computation of the third, fifth and seventh order coefficients (also termed primary, secondary and tertiary aberration coefficients). The table gives the computation of paracanonical coefficients and hence only proceeds as far as the tertiary "p" coefficients, since only the "p" coefficients are required for the description of the aberrations of the system, provided that the initial paracanonical coordinates are those appropriate to the particular conjugate planes (see Section 4). However, primary and secondary "q" coefficients are computed, as these are required in the iteration formulae for the tertiary "p" coefficients (see Section 5).

In M Section 81 are listed the iteration formulae for the tertiary coefficients. These formulae are applicable directly to systems containing aspherical surfaces, the asphericity of a surface being reflected in the value of the intrinsic coefficients (see M Section 65,67). Since it would have been an advantage for future work to have a computer programme which was applicable to any type of system, the possibility of using the iterative formulae was considered. However, to use these formulae efficiently, high speed storage for about 300 words were required. Now, in the DEUCE there are 402 words of high speed storage. Owing to the magnitude of the numbers which occur during the computation, floating point arithmetic has to be used. This is done by subroutine requiring two delay lines of instructions (64 words). Also, delay line 11 is used for all drum transfers so this cannot be used as a "permanent" store. Thus the available high speed storage is 306 words. Floating point arithmetic requires two stores per word, so that the direct computation of the tertiary coefficients cannot be easily accomplished. However, in the iteration formulae, 55 products of the primary intermediate coefficients are used frequently. If only these are stored, then the computation of the coefficients requires approximately 1300 arithmetic operations, after the required intrinsic coefficients have been computed.

Now, the table published in reference 4 is based on the method of condensed iteration for spherical surfaces only, presented in M Section 84. The simplification comes about as a result of the factorization of $\Delta\Delta$. For spherical surfaces

$$\Delta \underline{\underline{A}} = \underline{\underline{I}} \underline{\underline{D}} = \underline{\underline{E}} \underline{\underline{J}}, \text{ from M 60.3} \quad (8.1)$$

Let the true expansion of $\underline{\underline{D}}$ be

$$\underline{\underline{D}} = \sum_{n=1}^{\infty} \sum_{\mu=0}^n \sum_{\nu=0}^{\mu} \hat{g}_{\mu\nu}^{(n)} \xi_1^{n-\mu} \eta_1^{\mu-\nu} \zeta_1^{\nu}, \quad (8.2)$$

and the pseudo-expansion of $\underline{\underline{D}}$ be

$$\underline{\underline{D}} = \sum_{n=1}^{\infty} \sum_{\mu=0}^n \sum_{\nu=0}^{\mu} \hat{g}_{\mu\nu}^{(n)} \xi_1^{n-\mu} \eta_1^{\mu-\nu} \zeta_1^{\nu}. \quad (8.3)$$

From M 9.51, $\underline{\underline{I}}$ can be written

$$\underline{\underline{I}} = i_p \{ (\underline{\underline{Y}}_1 + q \underline{\underline{V}}_1) + \sum_{n=1}^{\infty} \sum_{\mu=0}^n \sum_{\nu=0}^{\mu} (G_{\mu\nu}^{(n)\dagger} \underline{\underline{Y}}_1 + \bar{G}_{\mu\nu}^{(n)\dagger} \underline{\underline{V}}_1) \xi_1^{n-\mu} \eta_1^{\mu-\nu} \zeta_1^{\nu} \} \quad (8.4)$$

where $q = i_q/i_p$ and

$$G_{\mu\nu}^{(n)\dagger} = q \left(G_{\mu\nu p}^{(n)} - G_{\mu\nu q}^{(n)} \right), \quad (8.5)$$

and similarly for $\bar{G}_{\mu\nu}^{(n)\dagger}$ (these symbols correspond to the

$G_{\mu\nu(i)}^{(n)}$, $\bar{G}_{\mu\nu(i)}^{(n)}$ in M 84.15).

$$\text{Then, writing } g^{\dagger} = i_p \hat{g} \quad (8.6)$$

it is easy to arrive at expressions of the type

$$s_i = s_i^{\dagger} + A^{\dagger} \underline{\underline{a}}, \quad (8.7)$$

$$\bar{s}_i = q s_i^{\dagger} + \bar{A}^{\dagger} \underline{\underline{a}}. \quad (8.8)$$

The complete list for the secondary contributions is given in M 84.21-22. For the tertiary contributions, one has expressions of the type

$$t_i = t_i^{\dagger} + A^{\dagger} s_i^{\dagger} + S_i^{\dagger} \underline{\underline{a}}, \text{ from M 84.23} \quad (8.9)$$

The formulae for obtaining the s^+ , t^+ are given by M 84.31, 84.33 in conjunction with (8.6). The primary intermediate q coefficients are obtained from the p coefficients by means of the identities given in M Section 20.

Having computed the G^+ , by introducing a new quantity \underline{G} such that

$$\underline{G} = \bar{G}^+ - qG^+ \quad (8.10)$$

the barred secondary contributions can be obtained from the unbarred contributions as illustrated by the following expression

$$\bar{s}_i - qs_i = \underline{A} \underline{a} \quad (8.11)$$

The complete set for the secondaries is given by M 84.42. From (8.5) one would expect that in order to calculate the \bar{S}^+ , the \bar{S}_q are required, however these are not computed in the table of reference . Having calculated the S_q by means of the identities M 21.6, the \bar{S}^+ may be obtained by means of recurrence relations. using quantities of the previous surface. These are indicated in the table by asterisks preceeding the quantities in question (see entries 102, 104, 106, ... 114 of the table). A typical recurrence relation is given on p.752 of reference 4.

Using this condensed iteration scheme, storage is required for about 45 terms, and approximately 525 operations are needed for the computation of the tertiary coefficients. Thus this method is much more suitable for use on a computer. However, as mentioned before, the scheme is applicable to systems containing spherical surfaces only.

The programme follows the steps in the published table except for two sets of results. As mentioned above, in the table the \bar{S}^+ are computed from results for the previous surface. Since this would involve a considerable increase in the number of results to be stored on the drum, for which the access time is rather long (up to 63 msec), it was decided to compute the \bar{S}^+ from the definitive equation

$$\bar{S}^+ = q\bar{S}_p - \bar{S}_q \quad (8.12)$$

For this the \bar{S}_q are required and these are not computed in the table. In the programme, \bar{S}_{1q} \bar{S}_{5q} are found using the identities of M Section 22 p.32.33. \bar{S}_{6q} is not given by any

identity and thus is computed using the last iterative formulae of M 11.3, which can be written as

$$\bar{S}_{6q} = 3c_q [q \{ \bar{C}_p + \frac{1}{g}(i_q^2 + i_q'^2 + v_q'^2 - 3v_q^2) \} - \bar{C}_q], \quad (8.13)$$

from which \bar{S}_{6q} is found by summation (3.16). A section of this expression, $\frac{1}{g}(i_q^2 + i_q'^2 + v_q'^2 - 3v_q^2)$, has the same form as the entry t_{34} of the table, the difference being that t_{34} used p coefficients, and the above, q coefficients. Hence this part of the expression is worked out by the same section of programme that computes t_{34} , but using q data, the result being stored until required in (8.13). Since \bar{S}_{6q} requires c_q , the $A_q \dots \bar{C}_q$ computed in the table in entries $t_{20} \dots t_{24}$, are, in the programme, obtained from $a_q \dots \bar{c}_q$, these being obtained from M20.21, 33, 34. This method also avoids retaining t_6 , t_{23} , t_{24} (q , C_q , \bar{C}_q) of the previous surface.

In the computer programme, as in the table, the parameters of the system are to be so arranged that, for the particular paraxial coordinates used, N_1/g is unity. Thus the antéprimes associated with the intermediate coefficients in the iteration formulae are omitted in the computations. The final aberration coefficients produced by the table are augmented, that is, they are the coefficients of the series (3.18). To obtain the displacements ε_k' , they have to be multiplied by $\mu (= 1/N_k v_{pk}')$. This is done by the programme before punching out, and, to assist in the comparison of magnitudes, the contributions to these coefficients by the surfaces are also multiplied by μ before punching. Telescopic systems cannot be handled by the programme since v_{pk}' is zero for these.

Using $\bar{O}T$ coordinates, with an object at infinity, the initial data for the paraxial traces are, $v_{a1} = 0$, $y_{a1} = 1$, $v_{b1} = 1$, $y_{b1} = p$ where p is the distance of the paraxial entrance pupil from the first polar tangent plane. If d is the distance of the diaphragm behind the (i) th surface then

$$p = \frac{y_{qi} - v_{qi}' d}{y_{pi} - v_{pi}' d} \quad (8.14)$$

the paraxial coefficients in this expression being canonical. To specify the system, the data given to the machine are 1) curvatures c_j ($j = 1 \dots k$); 2) separations d_j ($d_1 \equiv 0$);

3) ratio of refractive indices $k_j (=N_j/N_f)$; and 4) the refractive index of the image space N_f . With these data and those for the "a" and "b" paraxial traces, punched in floating decimal form, the machine proceeds through the system, one surface at a time. At each surface, all the augmented coefficients up to \bar{t}_{10pj} and the sums of the coefficients of the kind A_{pj} to \bar{T}_{10pj} are calculated before proceeding to the next surface. Note that coefficients $A_{pj} \dots$ are the intermediate coefficients of the $(j+1)$ th surface. At the last surface, the $A_{pk} \dots$ are the final augmented aberration coefficients of the entire system. The contributions to these by the surfaces, $a_{pj} \dots \bar{t}_{10pj}$ are, of course, stored at each surface to be punched out with the aberration coefficients. The paraxial coefficients are also stored for punching.

Thus the following results are punched out for each surface of the system: 1) paraxial coefficients $v_{pj}, y_{pj}, v_{dj}, y_{dj}$ (first order), 2) primary p coefficients (third order), 3) secondary p coefficients (fifth order), 4) tertiary p coefficients (seventh order); and finally 5) the third, fifth and seventh order coefficients of the entire system. The coefficients 2), 3), 4) and 5) are first multiplied by μ before punching. 1), 2), 3) and 4) constitute 41 coefficients per surface, being composed from 4 paraxial coefficients, 5 primary, 12 secondary, 20 tertiary. From the expansion of \bar{g}_k (3.18) 6 primary coefficients would be expected, but since $B = 2\bar{A}$, (M 20.32, 42) B is not punched.

Since it is impossible to assess the magnitude of the numbers occurring during computation, checks on the accuracy of the working cannot be easily performed. The only computation check used in this programme is as follows. The coefficients of the second surface of the system computed as an example in the table⁴ are stored in binary in the machine. At the end of the calculation of the coefficients of an entire system, the machine is redirected back into the programme, and is given the first surface results of the known system. It then computes the second surface results and compares these with the stored information in groups $t, \bar{t}, A, S, T, \bar{T}$. Should any one of these comparisons not agree, the punch-out routine will not be entered, and the particular failure will be indicated on the console. This method of checking in general only detects consistent failure of any one section of the computer. It is expected that should random errors occur, such as the insertion or loss of digits, then these will affect programme instructions, many of which pass through the arithmetic units. In this case, errors will be detected by virtue of the fact that the programme will fail.

The programme length is about 7,600 words, including result stores, and floating point subroutines are entered about 890 times for each surface. The computing time per surface is about 10 seconds, which gives an average time for obeying an instruction of about 400 μ sec. Since a particular subroutine must be entered at the same m.c. every time, this gives a minimum time between entries of 1 M.c. (≈ 1 m.sec) so, although the programme has been optimised as far as possible the subroutines place an upper limit on the minimum time possible to perform the computation in this programme. Including the punching out of the results, the machine time is approximately 45 seconds per surface.

A sample set of results is given in Appendix 1.

9. Ray Trace Programme

1. Theory

It is not feasible to obtain in closed form an expression for $\Delta\Delta$ as a function of the coordinates of a ray except in the case of a single surface (see M Section 15). Thus, in Section 3, $\Delta\Delta$ is expanded as a power series in Y_1, V_1 . The coefficients of the series thus characterize systems of rays rather than individual rays. However, for any particular ray, $\Delta\Delta_1, \Delta\Delta_2, \Delta\Delta_3, \dots$ may be calculated by ordinary algebraic means. Having found the $\Delta\Delta$'s, then \underline{e}' in the ideal image plane can be easily obtained by direct summation, if $\mu = 1$, by (3.12). In addition, the individual $\Delta\Delta$'s will give some information about the way the surfaces contribute to the final aberration for the particular ray. An examination of the $\Delta\Delta$'s for various principal rays for instance, can give some indication of which surfaces contribute mainly to distortion. This information can more readily be obtained from the contributions to the aberration coefficients, but if a ray trace is to be done, then any additional information apart from the final \underline{e}' is of advantage.

A ray trace scheme designed by Buchdahl on the basis of certain equations in M first computes at each surface the $\Delta\Delta$'s in (5.30) directly and then uses these in equations (5.29) as a means of tracing the ray through the system. This scheme not only has the advantage of obtaining the $\Delta\Delta$'s easily, but proves to be a very elegant way of tracing skew rays algebraically. The ray trace has, so far, been developed for system containing spherical surfaces only. Using paracanonical coordinates, the

theory is as follows.

From M 60.3, for spherical surfaces only,

$$\underline{\Delta\Lambda} = \underline{D} \underline{I} . \quad (9.1)$$

That is, writing in full

$$\begin{aligned} \Delta\Lambda_{ya} &= \underline{D}_a I_y , & \Delta\Lambda_{yb} &= \underline{D}_b I_y , \\ \Delta\Lambda_{za} &= \underline{D}_a I_z , & \Delta\Lambda_{zb} &= \underline{D}_b I_z , \end{aligned} \quad (9.2)$$

where

$$\underline{D} = \Delta N[y_0 (\underline{S} - 1) - v_0 x \underline{S}] , \quad (9.3)$$

$$\underline{I} = c \underline{Y} - \underline{V} , \quad (9.4)$$

and x is the x -coordinate of the point of intersection of the ray with the surface, measured from the pole of the surface.

\underline{S} in (9.3) is given by M 62.4, namely

$$\mu_2 \underline{S}^2 + 2\mu_1 \underline{S} + \mu_0 = 0 , \quad (9.5)$$

where

$$\begin{aligned} \mu_0 &= (k^2 - 1)(1 + \zeta)/k^2 \\ \mu_1 &= (1 - k)[c\eta - \zeta + (1 + \zeta)\underline{P}/k^2] \\ \mu_2 &= (1 - k)^2 [c^2 \xi - 2c\eta + \zeta - (1 + \zeta)\underline{P}^2/k^2] . \end{aligned} \quad (9.6)$$

In (9.6) a quantity \underline{P} appears. This is defined by M 56.72

$$\underline{P} = 1 - cx \quad (9.7)$$

whence

$$x = (1 - \underline{P})/c . \quad (9.8)$$

Now, for a spherical surface, if x, y, z are the coordinates of a point on the surface,

$$x^2 - 2rx + (y^2 + z^2) = 0. \quad (9.9)$$

Substituting x for \bar{x} in (1.6) we get

$$y^2 + z^2 = \xi - 2\eta x + \zeta x^2, \quad (9.10)$$

thus (9.9) becomes

$$(1 + \zeta)x^2 - 2x(r + \eta) + \xi = 0. \quad (9.11)$$

Using (9.8) and (9.11) one obtains an equation for \underline{P} , namely

$$(1 + \zeta)\underline{P}^2 + 2(c\eta - \zeta)\underline{P} + (c^2\xi - 2c\eta + \zeta - 1) = 0 \quad (9.12)$$

Let

$$p_2 = 1 + \zeta = 1 + V^2 + W^2,$$

$$p_1 = c\eta - \zeta = c(YV + ZW) - p_2 + 1$$

$$p_0 = c^2\xi - 2c\eta + \zeta - 1 = c^2(Y^2 + Z^2) - 2p_1 - p_2 \quad (9.13)$$

then

$$\underline{P} = [\sqrt{(p_1^2 - p_0 p_2)} - p_1]/p_2, \quad (9.14)$$

the positive root always being taken since $\underline{P} \rightarrow 1$ in the paraxial limit. Having found \underline{P} , then we are in a position to determine the μ_0, μ_1, μ_2 in (9.6).

First, however, it is convenient to consider, instead of \underline{S} a quantity

$$\bar{\underline{S}} = (1 - k)\underline{S}. \quad (9.15)$$

Then (9.5) becomes

$$\bar{\mu}_2 \bar{\underline{S}}^2 + 2\bar{\mu}_1 \bar{\underline{S}} + \bar{\mu}_0 = 0, \quad (9.16)$$

where

$$\bar{\mu}_2 = k^2 \mu_2 / (1 - k)^2, \bar{\mu}_1 = k^2 \mu_1 / (1 - k), \bar{\mu}_0 = k^2 \mu_0. \quad (9.17)$$

Thus

$$\bar{\mu}_0 = (k^2 - 1)(1 + \zeta) = (k^2 - 1)p_2$$

$$\bar{\mu}_1 = k^2(c\eta - \zeta) + (1 + \zeta)P = k^2 p_1 + p_2 \underline{P}$$

$$\bar{\mu}_2 = k^2(c^2\xi - 2c\eta + \zeta) - (1 + \zeta)\underline{P}^2 = k^2(p_0 + 1) - p_2 \underline{P}^2. \quad (9.18)$$

Then

$$\bar{\underline{S}} = [\pm \sqrt{(\bar{\mu}_1^2 - \bar{\mu}_0 \bar{\mu}_2) - \bar{\mu}_1}] / \bar{\mu}_2, \quad (9.19)$$

the positive root being taken when k is positive, the negative root when k is negative. As a result of (9.15), (9.3) becomes

$$\underline{D} = N'[y_0(\bar{\underline{S}} + k - 1) - v_1' x \bar{\underline{S}}]. \quad (9.20)$$

Let

$$\bar{Q} = \bar{\underline{S}} + k - 1,$$

then

$$\begin{aligned} \underline{D}_a &= N'[y_a \bar{Q} - v_a' x \bar{\underline{S}}], \\ \underline{D}_b &= N'[y_b \bar{Q} - v_b' x \bar{\underline{S}}], \end{aligned} \quad (9.21)$$

since

$$\underline{D} = \underline{D}_a s_0 + \underline{D}_b t_0.$$

The transfer from one surface to the next is accomplished using (5.29) and (5.30) as follows.

In this work, where S_y, S_z occur as subscripts, the letters will be written side by side, in order to simplify the printing.

Using paracanonical coordinates, (5.29) and (5.30) become

$$Y_j = y_{aj} (S_y + \delta_{syj}) + y_{bj} (T_y + \delta_{tyj})$$

$$Z_j = y_{aj} (S_z + \delta_{szj}) + y_{bj} (T_z + \delta_{tzj})$$

$$\begin{aligned}
V_j &= v_{aj} (S_y + \delta_{syj}) + v_{bj} (T_y + \delta_{tyj}) \\
W_j &= v_{aj} (S_z + \delta_{szj}) + v_{bj} (T_z + \delta_{tzj}), \quad (9.22)
\end{aligned}$$

where

$$\begin{aligned}
\delta_{syj} &= - \sum_{i=1}^{j-1} \Delta\Lambda_{sybi} \\
\delta_{szj} &= - \sum_{i=1}^{j-1} \Delta\Lambda_{szbi} \\
\delta_{tyj} &= + \sum_{i=1}^{j-1} \Delta\Lambda_{tyai} \\
\delta_{tzj} &= + \sum_{i=1}^{j-1} \Delta\Lambda_{tza i}, \quad (9.23)
\end{aligned}$$

the antepimes being omitted in (9.23), since N_i/g is arranged to be unity here, as in the computation for the aberration coefficients. To simplify the writing further, δ_{sy} shall be denoted by δ_{yb} , δ_{sz} by δ_{zb} ..., and $\Delta\Lambda_{syb}$ shall be denoted by $\Delta\Lambda_{yb}$,

Now, from (9.23), we get

$$\begin{aligned}
\delta_{yb(j+1)} &= \delta_{ybj} - \Delta\Lambda_{ybj} \\
\delta_{zb(j+1)} &= \delta_{zbj} - \Delta\Lambda_{zbj} \\
\delta_{ya(j+1)} &= \delta_{yaj} + \Delta\Lambda_{yaj} \\
\delta_{za(j+1)} &= \delta_{zaj} + \Delta\Lambda_{zaj} \quad (9.24)
\end{aligned}$$

By adding S_y , S_z ... to both sides of the appropriate equations (9.24), one obtains

$$\begin{aligned}
(S_y + \delta_{yb})_{j+1} &= [(S_y + \delta_{yb}) - \Delta\Lambda_{yb}]_j \\
(S_z + \delta_{zb})_{j+1} &= [(S_z + \delta_{zb}) - \Delta\Lambda_{zb}]_j \\
(T_y + \delta_{ya})_{j+1} &= [(T_y + \delta_{ya}) + \Delta\Lambda_{ya}]_j \\
(T_z + \delta_{za})_{j+1} &= [(T_z + \delta_{za}) + \Delta\Lambda_{za}]_j. \quad (9.25)
\end{aligned}$$

Thus, given the $\Delta\Lambda$'s and $(S_y + \delta_{yb}), \dots$ at j , Y_{j+1}, \dots can be found using (9.25) and (9.22). Note here that $(\underline{S} + \underline{\delta})_1 = \underline{S}$, and that $(\underline{S} + \underline{\delta})_j$ stands for $(\underline{S} + \underline{\delta}_j)$, since \underline{S} is defined at the first surface only. (9.25) are written as they are to indicate that $(\underline{S} + \underline{\delta})$ is to be evaluated at each surface for use in (9.2).

To summarize, the ray trace proceeds thus: given $(\underline{S} + \underline{\delta})$ at a particular surface, Y, Z, V, W are computed using (9.2). From these, the p_0, p_1, p_2 are obtained next, by (9.13) from which \underline{P} is calculated using (9.14). Then the $\bar{\mu}_0, \bar{\mu}_1, \bar{\mu}_2$ are obtained by (9.18) and thus \bar{S} from (9.19). Next \underline{D}_a and \underline{D}_b are computed by (9.21), and finally the $\Delta\Lambda$'s from (9.2) and (9.4). Then, using (9.25) the $(\underline{S} + \underline{\delta})$ of the next surface can be obtained. It will be noticed that this scheme is a skew ray trace scheme in the sense that tangential and sagittal rays are special cases determined only by the starting data, namely, $S_z = 0$ for tangential rays, $S_y = 0$ for sagittal rays.

Having computed the $\Delta\Lambda$'s at each surface, then the displacements of a ray from the ideal image point in the ideal image plane can be directly obtained by virtue of (3.12). For paracanonical coordinates, bearing in mind (4.17)

$$\begin{aligned} \epsilon_{yk} &= \mu \sum_{i=1}^k \Delta\Lambda_{yai} \\ \epsilon_{zk} &= \mu \sum_{i=1}^k \Delta\Lambda_{zai} \end{aligned} \quad (9.26)$$

From (9.24) and (9.25)

$$\begin{aligned} \sum_{i=1}^k \Delta\Lambda_{yai} &= (T_y + \delta_{ya})'_k - T_y \\ \sum_{i=1}^k \Delta\Lambda_{zai} &= (T_z + \delta_{za})'_k - T_z \end{aligned} \quad (9.27)$$

In practice, the object point is chosen to lie in the meridional plane, so that $T_z = 0$, hence

$$\sum_{i=1}^k \Delta\Lambda_{zai} = (T_z + \delta_{za})'_k \quad (9.28)$$

If the system is working in air, and has unit focal length, then $\mu = 1$, so that for this case

$$\begin{aligned}\varepsilon_{yk}' &= (T_y + \delta_{ya})_k' - T_y \\ \varepsilon_{zk}' &= (T_z + \delta_{za})_k' ,\end{aligned}\quad (9.29)$$

the quantities in brackets being computed in the image space by the scheme.

The refractive indices of a system are chosen appropriate to the particular wavelength (colour) of light. These values of N determine a particular value of v_{ak}' and hence the position of the ideal image plane, since the axial distance l_{ok}' of the image plane from the pole of the last surface is given by

$$l_{ok}' = y_{ak}' / v_{ak}' . \quad (9.30)$$

To trace rays of a different colour, a new set of values of N are required, which in turn fix the position of the new ideal image plane. Equations (9.29), which allow simple computation of ε_k' , always give ε_k' in the ideal image plane corresponding to the colour of the light. However, one usually requires that for various colours, all the displacements be obtained in one plane, whose position is determined by a particular "base" colour. In these circumstances, ε_k' can be obtained from M 86.1, namely

$$\begin{aligned}\varepsilon_{ky}' &= {}_o l_{ok}' V_k - Y_k - {}_o h_k' , \\ \varepsilon_{kz}' &= {}_o l_{ok}' W_k - Z_k' ,\end{aligned}\quad (9.31)$$

where

$$\begin{aligned}{}_o l_{ok}' &= {}_o y_{ak}' / {}_o v_{ak}' , \\ {}_o h_k' &= {}_o T_y / {}_o N_{k1} {}_o v_{ak}' .\end{aligned}\quad (9.32)$$

The preceding subscript 'o' refers to quantities associated with the base colour. The factor N_1 in ${}_o h_k'$ (see 4.21) has been omitted here since OT coordinates are considered to be the most

useful set to use for the ray trace. As mentioned before, $g = 1$ for this set, thus in order to be able to omit the antéprimes associated with the $\Delta\Delta$'s, N_1 has to be made unity if it is not so already.

If expressions (9.31), (9.32) are used to compute ε_k' , then it is possible to allow for any desired image plane shift.

For a new image plane F' which is at a distance x' from the ideal plane,

$$\tilde{l}_{ok}' = {}_o l_{ok}' + x' , \quad (9.33)$$

and for this plane,

$$\tilde{h}_k' = {}_o h_k' + x' {}_o v_{ok}' {}_o T_y . \quad (9.34)$$

Hence

$$\begin{aligned} \tilde{\varepsilon}_{ky}' &= \tilde{l}_{ok}' V_k' - Y_k' - \tilde{h}_k' , \\ \tilde{\varepsilon}_{kz}' &= \tilde{l}_{ok}' W_k' - Z_k' . \end{aligned} \quad (9.35)$$

There are several checks which can be used during the computation. One of these involves the calculation of an optical invariant αE_x , defined as

$$\begin{aligned} \alpha E_x &= \alpha N r I_x \\ &= \alpha N (ZV - YW) \\ &= N (ZV - YW) / \sqrt{1 + V^2 + W^2} \\ &= N (ZV - YW) / \sqrt{p_2} . \end{aligned} \quad (9.36)$$

As the factor $(ZV - YW)$ is identically zero for tangential rays the following checks can be used in this case

$$(1 - \underline{P})(V^2 + 1) = c[(YV + r) \mp \sqrt{r^2 + 2YVr - Y^2}] \text{ as } r \gtrless 0.$$

and, according as $c \gtrless 0$,

$$\begin{aligned} \underline{S} &= \pm c \sqrt{p_2} \left[\sqrt{\{r^2 - [Y + V'(d' - r)]^2 / p_2\}} - \right. \\ &\quad \left. - k \sqrt{\{r^2 - (Y - rV)^2 / p_2\}} \right] , \end{aligned} \quad (9.37)$$

where the subscript + refers to quantities at the next surface. If \bar{S}_k is being checked, then $d' = 0$, and the subscript + relates to quantities after refraction at the last surface. These checks are rather unwieldly and are not used in the ray trace programme since the machine is adequately checked when tracing skew rays, these rays being in the majority in most pencils. The checks (9.37) were designed for use with desk machine computations.

Equations (9.22) require the paraxial coefficients y_a , v_a , y_b , v_b . These are obtained from a pair of paraxial ray traces, the starting data for which is given by (4.12). For OT coordinates which are used in the ray trace, the constants in (4.12) are given by (4.16). The p required for these constants is given by (8.14).

It is convenient for the numerical calculation of the paraxial coefficients, both canonical and paracanoncal, to rewrite equations (2.5) in the following form.

$$\begin{aligned} v &= *v' \\ y &= *y - dv \\ i &= cy - v \\ v' &= (1 - k) i + v, \end{aligned} \quad (9.38)$$

where $*y_j = y_{j-1}$ and $*v' = v'_{j-1}$.

By way of example, a ray of a pencil from an object at infinity will be traced through a wide angle system. The system is the Pantoskop, Fig. 8 (Flügge, p. 283), the dimensions of which have been adjusted to give it unit focal length. The diaphragm is 0.061849 units behind the third surface.

Table I shows the paraxial "p" and "q" traces necessary to determine p , the position of the entrance pupil. To compute p it is only necessary to proceed as far as the third surface; the trace in the table has been continued to the 6th surface in order to check that the focal length is in fact unity ($f' = 1/v_{pk}'$).

From Table I and using (8.14),

$$p = 0.091520,$$

in the diagram of the system, this is the distance AE.

The ray chosen for the ray trace is one of the 45°

pencil whose S_y , S_z are 0.015625, 0.015625, S_y being shown in the diagram.

With the object at infinity and in the tangential plane, for $\bar{O}T$ coordinates

$$\sigma = 1, \bar{\sigma} = -0.091520, \tau = 0, \bar{\tau} = 1,$$

thus the starting data for the "a" and "b" paraxial traces are

$$a - \text{ray } y_1 = 1, v_1 = 0,$$

$$b - \text{ray } y_1 = 0.091520, v_1 = 1.$$

At the first surface, the initial data for the actual ray trace is of course

$$S_y = 0.015625$$

$$S_z = 0.015625$$

$$T_y = 1.0$$

$$T_z = 0.0$$

$$T_y = V_1, \text{ for an object at infinity } (V_1 = \tan 45^\circ).$$

Table II shows the ray trace computation.

From Table II one obtains

$$l'_{dk} = y'_{dk} / v'_{dk} = 0.908481, h'_k = T_y / v'_{dk} = 1.0$$

$$e'_{ky} = (T_y + \delta)_k - T_y = l'_{dk} V'_k - Y'_k - h'_k = 0.003189$$

$$e'_{kz} = (T_z + \delta)_k = l'_{dk} W'_k - Z'_k = 0.002148.$$

Suppose a new image plane is distant 0.015625 from the ideal image plane. Then

$$x' = 0.015625, \tilde{l}'_{dk} = 0.924106, \tilde{h}'_k = 1.015625$$

$$\tilde{e}'_{ky} = 0.003498, \tilde{e}'_{kz} = 0.002428.$$

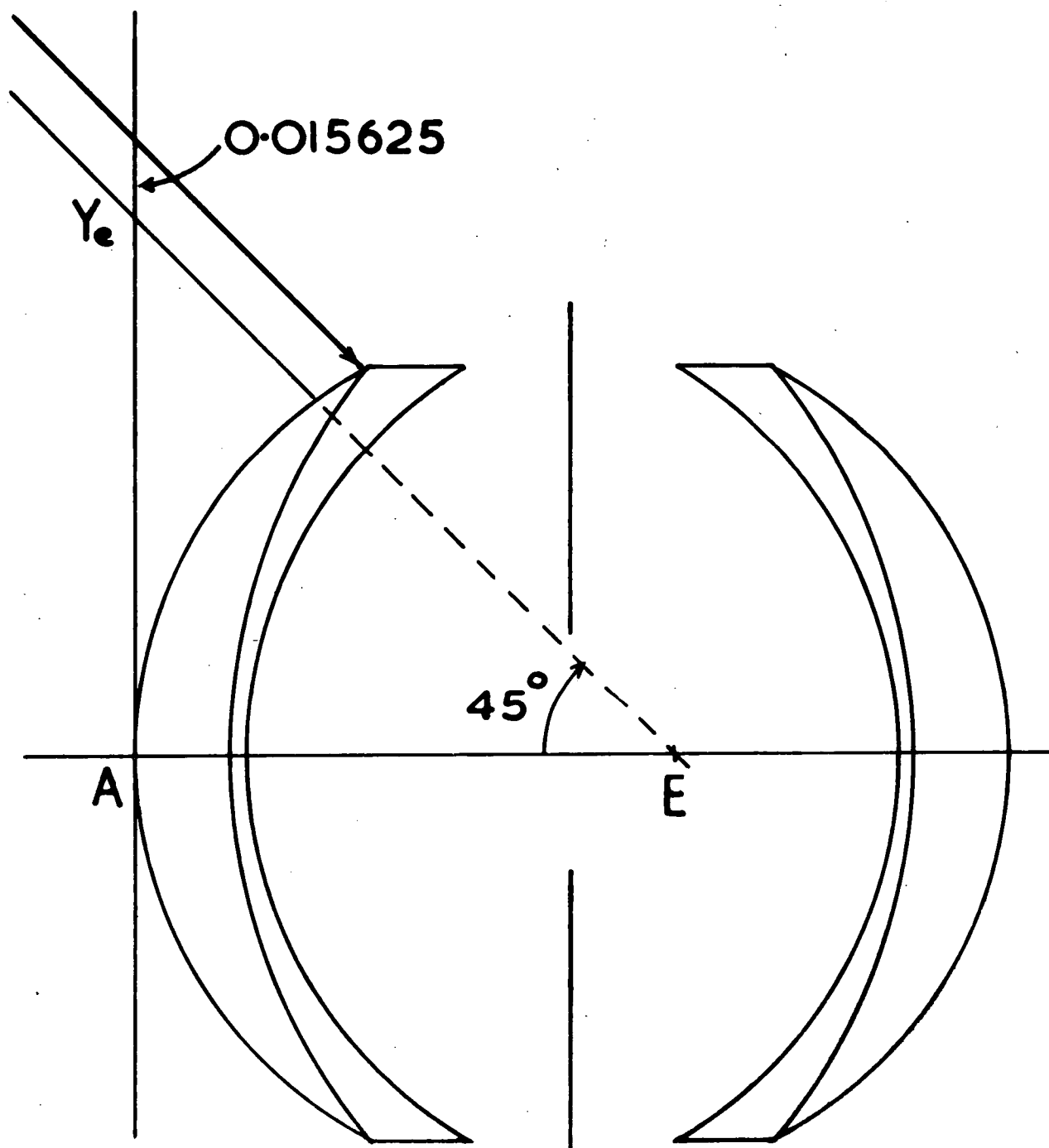


Fig. 8 Pantoskop

TABLE I

j		1	2	3	4	5	6
c		13.45565	9.648182	13.13822	-13.13822	-9.648182	-13.45565
d			0.017357	0.001895	0.123697	0.001895	0.017357
N		1.0	1.5331	1.6079	1.0	1.6079	1.5331
k		0.652273	0.953480	1.6079	0.621929	1.048790	1.5331
p ray	v_p	<u>0.0</u>	4.678890	4.873611	0.571964	-3.810739	-3.598413
	$*y-dv=y_p$	<u>1.0</u>	0.918786	0.909549	0.838798	0.846021	0.908481
	$cy-v=i_p$	13.45565	4.185726	7.076241	-11.59228	-4.351828	-8.62579
	$(1-k)i+v=v_p'$	4.678890	4.873611	0.571964	-3.810739	-3.598413	1.0
q ray	v_q	<u>1.0</u>	0.652273	0.616848	1.091591	1.411640	1.409813
	y_q	<u>0.0</u>	-0.011322	-0.012491	-0.147518	-0.150194	-0.174664
	i_q	-1.0	-0.761508	-0.780957	0.846533	0.037455	0.940410
	v_q'	0.652273	0.616848	1.091591	1.411640	1.409813	0.908480
$d=0.061849, p=-(y_{q3}-v_{q3}'d)/(y_{p3}-v_{p3}'d)=0.091520$							

TABLE II

j		1	2	3	4	5	6	k'
c		13.45565	9.648182	13.13822	-13.13822	-9.648182	-13.45565	
d			0.017357	0.001895	0.123697	0.001895	0.017357	
N'		1.5331	1.6079	1.0	1.6079	1.5331	1.0	
k		0.652273	0.953480	1.60790	0.621929	1.048790	1.5331	
y_a	t_1	<u>1.0</u>	0.918786	0.909549	0.838798	0.846021	0.908481	
v_a	t_2	<u>0.0</u>	4.678890	4.873611	0.571964	-3.810739	-3.598413	<u>1.0</u>
y_b	t_3	<u>0.091520</u>	0.072766	0.070751	-0.070751	-0.072766	-0.091520	
v_b	t_4	<u>1.0</u>	1.080486	1.062881	1.143938	1.062881	1.080486	1.0
$^*(t_5 - t_{29}) = (S_y + \delta)$	t_5	0.015625	0.018790	0.018872	0.018930	0.016044	0.016608	0.016578
$^*(t_6 - t_{30}) = (S_z + \delta)$	t_6	0.015625	0.017131	0.017117	0.016438	0.015211	0.015284	0.015767
$^*(t_7 + t_{27}) = (T_y + \delta)$	t_7	1.0	1.084969	1.075404	1.078590	0.993350	1.005195	<u>1.003189</u>
$^*(t_8 + t_{28}) = (T_z + \delta)$	t_8	0.0	0.040444	0.042069	0.004702	-0.031529	-0.030013	<u>0.002148</u>
$t_1 t_5 + t_3 t_7 = Y$	t_9	0.107145	0.096212	0.093251	-0.060433	-0.058708	-0.076908	<u>-0.076751</u>
$t_1 t_6 + t_3 t_8 = Z$	t_{10}	0.015625	0.018683	0.018546	0.013456	0.015163	0.016632	<u>0.014127</u>
$t_2 t_5 + t_4 t_7 = V$	t_{11}	1.0	1.260209	1.235003	1.244667	0.994674	1.026336	<u>1.019767</u>
$t_2 t_6 + t_4 t_8 = W$	t_{12}	0.0	0.123855	0.128138	0.014780	-0.091479	-0.087426	<u>0.017915</u>
$t_{11}^2 + t_{12}^2 + 1 = P_2$	t_{13}	2.0	2.603468	2.541651	2.549415	1.997744	2.061009	2.040245
$c(t_9 t_{11} + t_{10} t_{12}) - t_8 + 1 = P_1$	t_{14}	0.441708	-0.411322	0.002641	-0.563787	-0.420950	0.020651	
$c^2(t_9^2 + t_{10}^2) - 2t_{14} t_3 = P_0$	t_{15}	-0.760691	-0.886638	-0.986559	-0.760184	-0.813600	-0.981332	
$t_{14}^2 - t_{13} t_{15}$	t_{16}	1.716488	2.477519	2.507495	2.255881	1.802564	2.022961	
$(\sqrt{t_{16} - t_{14}})/t_{13} = P$	t_{17}	0.434220	0.762573	0.621984	0.810282	0.882769	-0.680083	

[illegible]

(2) Programme

The programme constructed on the basis of this scheme will operate on systems containing 15 surfaces or less. The ray trace scheme presented above is very suitable for machine computation. Owing to the fact that there are no axial intersection lengths of rays to be computed, the numbers appearing during the course of the calculation are of modest size. There are no trigonometric functions to be evaluated, and only two square roots are required per surface, or three, including the αE_x check. Working to an accuracy of six decimal places with fixed point operation, the upper limit of the DEUCE machine is 2^7 . This restricts the initial value of T_y to be less than 64, which corresponds to an angle of about 80° for a pencil from an infinite object. Note here that the ray trace scheme will not work for angles $>90^\circ$ in any case. Thus fixed point working restricts the machine to systems of 160° full field or less, which is adequate for most of the systems generally considered.

It is desirable to work in fixed point arithmetic since, in computing a spot diagram, many rays are required and in these cases rapid computation is of considerable economic importance. In view of the need for fast calculation, considerable effort was directed towards optimization of the programme, to the extent that special subroutines were designed to carry out the multiplication and divisions for the particular number size used throughout the computation. The machine time per surface is 0.5 seconds, 50 per cent longer than the minimum time in which the instructions could be obeyed if all the stores were immediately accessible.

Also in the design of the programme, thought was given to the question of making it as universal in its application as possible. Accordingly, the present programme can produce automatically the $\underline{e_k'}$ for any one of four types of trace, 1) a full pencil, 2) a tangential fan, 3) a sagittal fan, 4) a single specified ray. These results can be obtained for any pair of object and image planes, not necessarily conjugate, providing of course that only finite images are contemplated. The punch-out programme restricts the displacement $\underline{e_k'}$ to be less than the focal length (unity) but this could be altered should circumstances warrant it. The object point is always considered to be in the meridional plane, the only restriction on the object height is that T_y be less than 64. Note here that if V ,

corresponds to the angle of the principal ray of any pencil, then

$$T_y = H_{y1}/l_{o1} = V_1/v_{b1} . \quad (9.39)$$

In the programme, any particular pencil is characterized by its V_1 , the object height as such does not enter into the calculation. Knowing V_1 , l_{o1} and p , the designer can compute H_{y1} if desired from

$$H_{y1} = V_1 (l_{o1} - p) . \quad (9.40)$$

The ideal image point in an out-of-focus image plane is defined as the intersection point with the plane of the principal ray of the pencil when its path through the system is determined from the laws of paraxial optics. From this definition (9.34) can be deduced.

In addition, if the refractive indices corresponding to different coloured light for a given system is fed in subsequent to tracing rays through the base colour system, then any of the four sets of results can be obtained for the new colour. In this case, as before, the image plane can be at any location. However, to establish the image height in the base ideal image plane for use in tracing "coloured" rays, at least the principal ray of the required "coloured" pencils must also be traced through the base system.

To enable the machine to trace automatically the rays for a full pencil, it was necessary to design a vignetting scheme. This scheme causes the machine to cease computation on any ray whose intersection point with any surface and with the plane of the diaphragm lies outside the limits set by the designer. The individual limits for each surface and the diaphragm are independent of each other. In the work done in this thesis, the actual rims of the components were used so that all rays traced will in fact go through the actual system as constructed. Should tracing of rays be required which are vignettted by a particular surface alone, this is achieved by simply making all other limits so large that they cannot vignette before the selected one does. When the actual rims of the components do the vignetting, then, as will be shown later, the shape of the entrance pupil for the particular pencil can be obtained from the output results.

The programme requires as its initial data the curvatures (c) separations (d) and refractive indices (N) of the surfaces, V_{a1} , V_{a2} , V_{b1} , V_{b2} for the "a" and "b" trace, and for

vignetting purposes, the radii (ρ) of the rims of the components (supposed or actual) and the radius (ρ_d) and location (D) of the diaphragm. As mentioned before, the y , v for the "a" and "b" trace contain in them the position of the object and the paraxial entrance pupil. Since "p" and "q" traces are carried out during the initial design of a system, the entrance pupil position p is supposed to be known by the time ray traces are required. Thus nothing is gained by programming the machine to compute the "a" and "b" starting data by performing a "p" and "q" trace of its own.

From the initial data, the machine computes the "a" and "b" paraxial coefficients of the system, which are stored until all the ray traces for the particular system are completed. Then the machine calls for a "ray data card" on which is punched, in binary, 1) V_1 of the principal ray, 2) S_y , S_z , the value of these depending on the type of trace required, 3) the "scan interval" δ . From V_1 , and the paraxial data, αT_y , αl_k , αh_k are computed, which are stored initially for the duration of the particular trace. Next, the coordinates S_y , S_z from the card are taken as the starting data for the ray trace. For the first three types of trace, S_y and S_z must be zero on the ray data card. The machine then proceeds to trace the ray in accordance with the scheme in table II, also computing and comparing αE_x , as it goes. At each surface also, vignetting is examined.

Now, for any spherical surface, equation (9.9) holds.

This can be written

$$c^2(y^2 + z^2) + \left(\frac{r-x}{r}\right)^2 - 1 = 0$$

$$\text{or } c^2(y^2 + z^2) + \underline{P}^2 - 1 = 0, \text{ using (9.7).}$$

If $y^2 + z^2 \leq \rho^2$ for a particular surface, the vignetting will not occur. In this case then,

$$c^2\rho^2 + \underline{P}^2 - 1 \geq 0 \quad (9.41)$$

For a plane surface, the corresponding condition is

$$\rho^2 \geq Y^2 + Z^2. \quad (9.42)$$

Depending on the value of c , either (9.41) or (9.42) is used to determine whether the ray is vignetted.

To determine whether the diaphragm vignettes, the coordinates of the ray in the plane of the diaphragm are found from (1.6), where \bar{x} here is the distance of the diaphragm from the following surface. Then of course the ray is not vignettted if $\rho_d^2 \geq \bar{y}^2 + \bar{z}^2$.

Assuming that the ray is not vignettted anywhere, when the image space is reached, \underline{e}_k' is computed by means of (9.35), taking into account the value of x' given on the ray data card. Then a pair of digits associated with V_i are examined to determine which of the four types of trace is required. If a full pencil is indicated, then the initial S_y , S_z will have been zero, and thus the principal ray will have been traced first. The machine will then return to the beginning of the ray trace programme and add $\bar{\delta}$ to the value of S_z and start again. Thus the next ray traced will have coordinates, $0, \bar{\delta}$. On completion of this ray, $\bar{\delta}$ is again added to S_z and the trace entered again. In this manner, a sagittal fan is traced, the \underline{e}_k' from each ray being stored. Eventually, vignetting will occur somewhere in the system. At this point the trace is immediately terminated, S_z is made zero and $\bar{\delta}$ added to S_y . In this manner the first quadrant of the entrance pupil is scanned by series of rays spaced at intervals of $\bar{\delta}$ horizontally and vertically. When a ray is encountered whose S_z is zero, and is vignettted by the system, the first quadrant scan is terminated. If, in addition, V_i is zero this completes the scanning for the axial pencil, since all four quadrants are symmetrical in this case. If V_i is not zero, then the lower quadrant of the entrance pupil is scanned, the first ray traced in this group having coordinates $-\bar{\delta}, 0$. As before, this quadrant is fully scanned when a ray is encountered whose S_z is zero and which won't go through the system, Fig. 9. Each time a ray is traced or vignettted the pair of digits characteristic of the type of trace is examined and the behaviour of the machine is varied according to the type of trace required. The tangential and sagittal fans are simple variations of the full pencil procedure.

In order to indicate the shape of the entrance pupil and to specify the location of each ray traced in the "grid", along with the \underline{e}_k' for each ray, $S_y/\bar{\delta}$, $S_z/\bar{\delta}$ are stored, giving the Cartesian coordinates of the ray relative to the principal ray as origin with $\bar{\delta}$ as the unit of length. If, in the example of table II, $\bar{\delta} = 2^{-3}$, then the (y, z) of this ray would be $(+4, +4)$. From these figures for a full pencil, the shape of the entrance pupil can be plotted. In addition, with each set of \underline{e}_k' punched out, V_k' , W_k' are also given, in case the designer wishes to compute X_s , X_t . Prior to punching any one set of results, in the absence of a particular

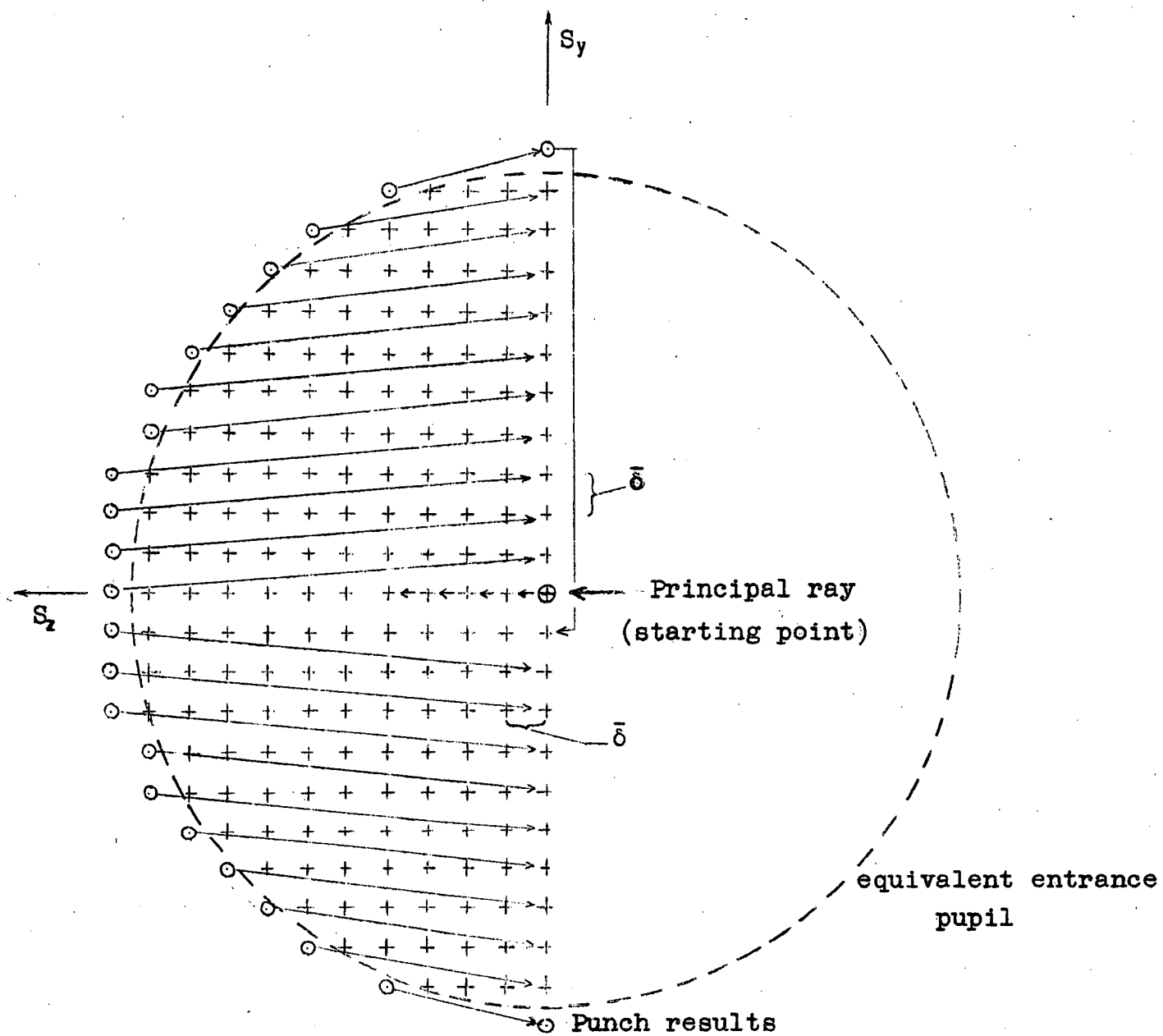


Fig. 9

Scan Procedure

- + Traced rays
- o Vignetted rays

binary digit on the ray data card, the T_y and h_k' of the pencil are stored as a pair for use in "colour tracing."

After a set of results is punched, the machine calls in a new ray data card, which can contain entirely new information i.e., new V_1 , different type of trace, new δ and x' . In this way the ray trace remains entirely under the control of the designer.

To trace rays of a different colour through the system, additional data are presented to the machine giving the new refractive indices. Included in these data is a parameter which tells the machine that the l_{0k}' is not to be computed from the paraxial coefficients but to be taken from the store, i.e. l_{0k}' . On the ray data cards used for "colour tracing" an extra digit is present which causes the machine to search the "T store" until it finds a T_y which matches the T_y computed from the V_1 on the card. The h_k' which is associated with this T_y is then extracted from the store and is subsequently used in the computation of ϵ_k' .

The computation of αE_x at each surface adequately checks the machine arithmetic for skew rays. To aid in the checking of tangential rays, the following relations are examined at each surface.

$$1) \quad \underline{P} < 1, \quad 2) \quad |\underline{S}| > |1 - k|, \quad 3) \quad t_0 \geq 0, \quad 4) \quad t_2 \geq 0 \\ \text{(see Table II).}$$

If the first two relations do not hold, then the machine has made a mistake. Relation 3) will not be true when the ray has missed the surface, in the sense that $y^2 + z^2 > r^2$, where r is the radius of the curved surface. If the ray is totally internally reflected at any surface, then $t_2 < 0$.

10. Predicted Displacements from Aberration Coefficients

In the theory of aberration coefficients, ϵ_k' is represented as an infinite series of homogeneous polynomials. If this series is terminated after the seventh order terms, then only an approximation to ϵ_k' can be obtained. This approximation will be denoted by $\epsilon'[3]$. $\epsilon'[3]$ is computed using (3.18) which, written in full, and omitting primes and subscripts, becomes

$$\begin{aligned}
\varepsilon'_y[3] = & (A\xi_1 + 2\bar{A}\eta_1 + \alpha_1 + S_1\xi_1^2 + S_2\xi_1\eta_1 + S_3\xi_1\zeta_1 + \\
& S_4\eta_1^2 + S_5\eta_1\zeta_1 + S_6\zeta_1^2)Y_1 \\
& + (T_1\xi_1^3 + T_2\xi_1^2\eta_1 + T_3\xi_1^2\zeta_1 + T_4\xi_1\eta_1^2 + T_5\xi_1\eta_1\zeta_1 + T_6\xi_1\zeta_1^2 + \\
& T_7\eta_1^3 + T_8\eta_1^2\zeta_1 + T_9\eta_1\zeta_1^2 + T_{10}\zeta_1^3)Y_1 \\
& + (\bar{A}\xi_1 + \bar{B}\eta_1 + \bar{C}\zeta_1 + \bar{S}_1\xi_1^2 + \bar{S}_2\xi_1\eta_1 + \bar{S}_3\xi_1\zeta_1 + \bar{S}_4\eta_1^2 + \\
& \bar{S}_5\eta_1\zeta_1 + \bar{S}_6\zeta_1^2)V_1 \\
& + (\bar{T}_1\xi_1^3 + \bar{T}_2\xi_1^2\eta_1 + \bar{T}_3\xi_1^2\zeta_1 + \bar{T}_4\xi_1\eta_1^2 + \bar{T}_5\xi_1\eta_1\zeta_1 + \bar{T}_6\xi_1\zeta_1^2 + \\
& \bar{T}_7\eta_1^3 + \bar{T}_8\eta_1^2\zeta_1 + \bar{T}_9\eta_1\zeta_1^2 + \bar{T}_{10}\zeta_1^3)V_1
\end{aligned} \tag{10.1}$$

$$\begin{aligned}
\varepsilon'_z[3] = & (A\xi_1 + 2\bar{A}\eta_1 + C\zeta_1 + S_1\xi_1^2 + S_2\xi_1\eta_1 + S_3\xi_1\zeta_1 + \\
& S_4\eta_1^2 + S_5\eta_1\zeta_1 + S_6\zeta_1^2)Z_1 \\
& + (T_1\xi_1^3 + T_2\xi_1^2\eta_1 + T_3\xi_1^2\zeta_1 + T_4\xi_1\eta_1^2 + T_5\xi_1\eta_1\zeta_1 + T_6\xi_1\zeta_1^2 + \\
& T_7\eta_1^3 + T_8\eta_1^2\zeta_1 + T_9\eta_1\zeta_1^2 + T_{10}\zeta_1^3)Z_1
\end{aligned} \tag{10.2}$$

In these expressions, the identity M20.42 has been used, namely $B = 2\bar{A}$. μ has been omitted from these expressions since the "coefficients programme" produces the aberration coefficients already multiplied by μ . Also, since paracanoncal $\bar{O}T$ coordinates are used throughout, the term involving W_1 (T_2) is absent in (10.2) as a result of the object points always being considered in the meridional plane. The relations (4.22) are to be kept in mind here.

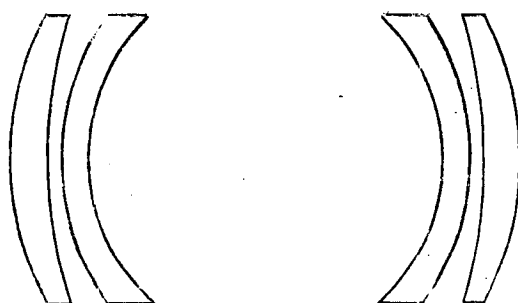
As well as the final aberration coefficients, this programme requires the paraxial coefficients of the surfaces of the system, all of these being part of the output from the coefficients programme. In addition, the approximate radii ρ of the components and the radius ρ_d and position D of the diaphragm are required. The paraxial coefficients and the ρ are used in a vignetting scheme in such a manner that this programme will produce the four types of results mentioned in the ray trace programme. However, here, vignetting cannot be based on the path of the actual ray for which $\varepsilon'_y[3]$ is computed since there is no way of determining its path using only the final aberration coefficients. This is not because the series for ε'_y has been terminated, but because in

PART III - Numerical Results

11. Specifications of systems examined.

In the specifications given for the following systems the dimensions have been so adjusted to give a focal length of unity. The systems are described by the curvatures c , the separations d and the refractive indices N . After these are given the distance D of the diaphragm from the first surface, the stop number ν of the full aperture as determined from an axial ray trace, and the approximate field of the system. The diagrams are not to scale.

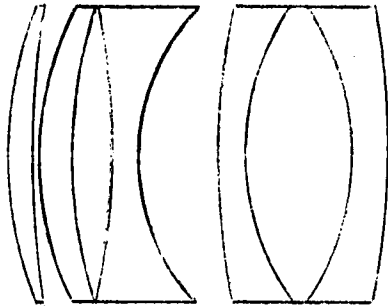
System 1 Vega



	c	d	N		
1	6.38090	0.0	1.0	D	0.370656
2	4.33692	0.061881	1.6116	ν	6.1
3	7.89645	0.000287	1.0	field \approx	90°
4	9.76526	0.006993	1.7480		
5	-9.76526	0.180758	1.0		
6	-7.89645	0.006993	1.7480		
7	-4.33692	0.0002874	1.0		
8	-6.38090	0.061881	1.6116		

64.

System 2 Sonnar



	c	d	N	
1	1.51831	0.0	1.0	D 0.429197
2	0.204230	0.093077	1.6515	v 1.6
3	2.83762	0.001263	1.0	field = 45°
4	1.16451	0.117215	1.6515	
5	-0.193567	0.070289	1.5095	
6	4.24506	0.018600	1.7467	
7	0.096097	0.156704	1.0	
8	1.64368	0.024988	1.5202	
9	-4.60076	0.198942	1.6515	
10	-0.960177	0.052451	1.5960	

equations (5.29), from which the path could be determined, the intermediate p and q coefficients are required. Only the p intermediate coefficients are produced by the coefficients programme. However, in order to prevent the machine from performing a scan indefinitely, at each surface Y_j, Z_j are computed using (5.32), that is, the increments δ (5.30) are neglected. Then the relation $Y_j^2 + Z_j^2 < \rho_j^2$ is examined and the ray is considered not vignetted if this holds. This method of vignetting is inaccurate as a result of two effects. Firstly, Y_j, Z_j are measured in the polar tangent plane, whereas ρ is supposed to represent the rim of the component. Thus, in the case of a "steep" surface, an actual ray approaching the rim of the surface and making a large angle with the axis may have $Y^2 + Z^2 > \rho^2$ and still pass through unvignetted. Secondly, since the δ 's have been omitted from the computation of the Y, Z , the value of Y_j, Z_j so obtained may be markedly in error if the system possesses large aberrations, or even large surface contributions and small final aberrations.

This programme was designed so that the displacements predicted by the first three orders of coefficients could be compared with those computed by ray traces. Thus, in all systems a set of accurately vignetted ray traces are available. By making p and ρ_j in the present programme sufficiently large, the pencil "traced" can always be so arranged that more than the required number of "rays" are computed. The correct set is then determined by comparison with the (y, z) coordinates in the ray trace results since here also, a set of (y, z) coordinates is produced with each "ray" computed. Computing more results than required tends to be wasteful of machine time. However, since the machine time per "ray" is 0.5 seconds, this is not serious. The machine time is almost independent of the number of surfaces in the system since the main body of the computation is taken up in the calculation of $e'[3]$, which of course is a fixed length operation. The vignetting section of the programme, which involves the surfaces individually is quite fast, and is only a small percentage of the total time.

Owing to the sizes of the numbers encountered in the coefficients of a reasonably well corrected system, it is possible to perform the computation to the end of the secondary coefficients in fixed point arithmetic with an accuracy of six decimal places. The tertiary section is done in floating point arithmetic. The results from the primary and secondary coefficients $e'[2]$ are

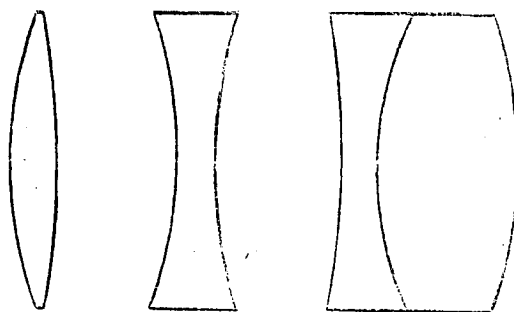
stored and punched with those from the entire set of coefficients so that the improvement in the accuracy of the predicted displacements due to the addition of the tertiary coefficients may be examined. Naturally a "ray" is examined for vignetting right through the system before any computation on coefficients is initiated. The results produced by this programme are $\epsilon'_y[2]$, $\epsilon'_z[2]$, $\epsilon'_y[3]$, $\epsilon'_z[3]$, y , z , the (y, z) of a "ray" being computed as in the ray trace programme. To simplify the tabulation the format of the results is the same as that of the ray trace programme.

The same ray data cards used in the ray trace are also used here. However, in this case, image plane shifts and "colour" cannot be handled. This is because the aberration coefficients give ϵ'_k only in the ideal image plane as determined by the value of y'_{ak}/v'_{ak} . The ideal image height does not directly enter into the power series expansion of ϵ'_k . In an out-of-focus plane, $\tilde{\epsilon}'_k$ can be computed from M36.4 namely

$$\tilde{\epsilon}'_k = \epsilon'_k + x' v'_{ak} \underline{S} + x' (\tilde{\epsilon}'_{ak} v'_{bk} - \tilde{\epsilon}'_{bk} v'_{ak}) \quad (10.3)$$

which, however, requires a complete set of a and b coefficients, and the computation of these is usually not worth the effort.

The object plane location must also be the same as that used in the calculation of the coefficients, owing to the fact that the "a" coefficients alone will describe the aberrations of a system only when the paracanonial coordinates are chosen appropriate to the object and entrance pupil positions, see Section 4.

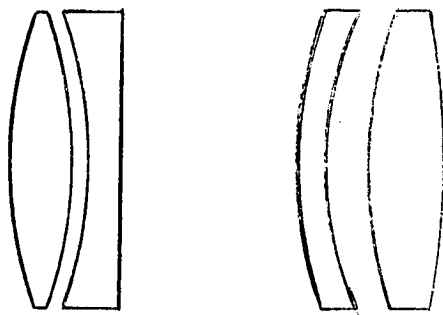
System 3 Tessar

c

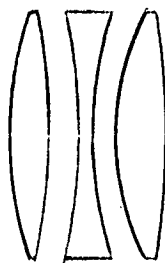
d

N

1	3.35531	0.0	1.0	D	0.1332
2	-0.063011	0.0350	1.6226	v	1.6
3	-2.16390	0.049839	1.0	field	= 50°
4	3.77245	0.0160	1.5761		
5	-0.651018	0.044260	1.0		
6	4.52661	0.01650	1.5282		
7	-3.08690	0.0500	1.6133		

System 4 Petzval Objective

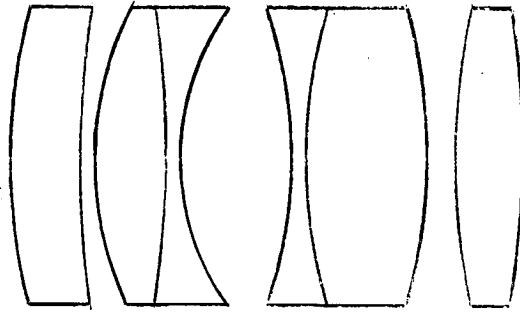
	c	d	N		
1	1.71310	0.0	1.0	D	0.239206
2	-1.44057	0.080301	1.5095	v	2.8
3	-1.57062	0.010337	1.0	field	= 20°
4	0.0	0.028674	1.61704		
5	1.48712	0.373180	1.0		
6	2.50710	0.024119	1.61704		
7	1.67694	0.039047	1.0		
8	-0.567497	0.054204	1.5095		

System 5 Telephoto

The specifications of this system are not available for publication.

ν 7.0

field $\approx 14^\circ$

System 6 Biotar

	c	d	N
1	1.19617	0.0	1.0
2	0.311526	0.10800	1.6424
3	2.23214	0.01700	1.0
4	-0.086957	0.15600	1.6231
5	3.53357	0.05100	1.5757
6	-2.59740	0.18900	1.0
7	1.98020	0.05100	1.6727
8	-1.87970	0.21200	1.6424
9	0.943396	0.01000	1.0
10	-0.833333	0.13900	1.6424

D 0.4210

 ν 1.6field $\approx 36^\circ$

Table III

	Semi-field angles of pencils	$\bar{\delta}$ ($f' = 1$)
Vega	0 10 20 30 35 40 45°	0.011719
Sonnar	0 5 10 15 20 25°	0.033203
Tessar	0 5 10 15 20 25°	0.011719
Petzval	0 2 4 6 8 10°	0.019531
Telephoto	0 1 2 3 4 5 6 7°	0.007813
Biotar	0 3 6 9 12 15 18°	0.039063

Table IV

	Vega	Sonnar	Tessar
A	- 0.7633	0.1778	0.8919
\bar{A}	- 0.1881	0.0242	0.1451
\bar{B}	0.1165	- 0.0727	- 0.0430
C	0.0342	0.1205	0.1017
\bar{C}	0.0270	0.2838	0.0048
S_1	- 661.2	- 1.389	- 77.24
\bar{S}_1	10.69	- 0.4599	- 4.923
S_2	43.10	- 1.756	- 19.31
\bar{S}_2	24.46	- 1.062	1.063
S_3	- 6.464	- 1.302	- 3.602
\bar{S}_3	- 0.2767	0.0728	- 0.6016
S_4	25.26	- 1.328	0.2279
\bar{S}_4	- 0.2515	- 1.063	- 1.646
S_5	- 0.1506	0.0898	- 1.480
\bar{S}_5	- 0.3547	- 0.0809	- 0.2835
S_6	- 0.1493	- 0.3708	- 0.5708
\bar{S}_6	- 0.0145	- 0.2756	- 0.0194
T_1	- 30473	- 8.233	- 2044
\bar{T}_1	- 213.1	- 1.953	- 214.1
T_2	1652	- 12.63	- 1337
\bar{T}_2	1023	- 2.384	42.47
T_3	- 884.8	- 7.220	- 146.9
\bar{T}_3	13.68	0.6395	- 19.97
T_4	3426	- 0.6876	238.1
\bar{T}_4	- 15.42	1.434	- 36.07
T_5	9.624	4.263	- 63.44
\bar{T}_5	- 9.484	1.896	- 13.68
T_6	- 9.070	- 0.8988	- 4.067
\bar{T}_6	- 0.1936	0.4851	- 1.607
T_7	- 40.05	2.547	- 10.72
\bar{T}_7	- 38.09	2.481	3.060
T_8	- 29.35	5.994	- 11.23
\bar{T}_8	- 1.000	0.3611	- 4.565
T_9	- 0.3155	0.6720	- 1.992
\bar{T}_9	0.4491	1.403	1.107
T_{10}	0.2097	- 0.6713	0.0271
\bar{T}_{10}	0.0159	- 0.3938	0.0764

Table IV (continued)

	Petzval	Telephoto	Biotar
A	0.0734	0.0862	0.0873
\bar{A}	0.0192	0.0559	- 0.0267
\bar{B}	- 0.1559	0.2026	- 0.1251
C	0.3388	0.0909	0.0761
\bar{C}	0.0903	1.710	- 0.2624
S_1	- 3.376	- 13.15	- 0.3541
\bar{S}_1	- 1.160	- 4.293	0.0606
S_2	- 4.556	- 17.19	0.2914
\bar{S}_2	- 1.975	- 12.02	- 1.811
S_3	- 1.413	- 6.873	- 1.512
\bar{S}_3	0.5900	- 0.3296	- 0.1433
S_4	- 2.033	- 12.39	- 1.864
\bar{S}_4	0.6692	- 7.386	0.5468
S_5	1.311	- 1.118	- 0.2593
\bar{S}_5	- 2.520	- 9.043	1.136
S_6	- 0.8556	- 3.405	- 0.3394
\bar{S}_6	0.3925	7.397	- 0.3351
T_1	- 29.17	- 102.7	- 1.552
\bar{T}_1	- 6.427	- 49.70	0.3322
T_2	- 43.61	- 289.2	1.901
\bar{T}_2	- 11.43	- 263.7	- 9.438
T_3	- 12.19	- 125.1	- 5.202
\bar{T}_3	- 0.3509	- 72.98	0.1982
T_4	- 19.43	- 410.6	-18.28
\bar{T}_4	- 3.790	- 332.9	- 0.2026
T_5	1.563	- 217.0	- 0.5094
\bar{T}_5	- 8.258	- 213.8	3.462
T_6	- 3.406	- 50.08	- 1.156
\bar{T}_6	4.390	- 0.4926	0.3703
T_7	- 0.0910	- 171.2	- 0.6656
\bar{T}_7	2.106	- 104.1	5.156
T_8	- 2.847	- 131.9	5.386
\bar{T}_8	10.07	- 90.79	1.488
T_9	5.081	- 3.832	0.8090
\bar{T}_9	- 7.456	- 93.91	1.686
T_{10}	- 1.441	- 19.37	0.3795
\bar{T}_{10}	1.495	43.18	- 0.6475

12. Calculations Performed on the Optical Systems.

1) Ray traces.

For an object plane at infinity, sets of ray traces were computed at several semi-field angles for each system, with the aberrations ϵ_k being calculated in the ideal image planes. The grid spacing δ for the full pencil scans was chosen so that at least 200 points were available in the axial spot diagram of each system. The various semi-field angles and the δ for each system are listed in Table III. The values of δ in this table appear to be complicated decimal numbers. This is because they are rounded-off decimal equivalents of the simple combinations of powers of two which were used on the ray data cards. It is easier to punch the binary digits for 2^{-6} and 2^{-7} on the cards in the case of the Biotar, for instance, than the binary representation of 0.04. ($2^{-5} + 2^{-7}$) is a sufficiently close approximation to ensure an adequate number of rays per pencil.

2) Aberration Coefficients.

For each of the six systems, the third, fifth and seventh order coefficients and the contributions thereto were calculated for an object plane at infinity. The final aberration coefficients are set out in Table IV, in which the primes and subscripts k have been omitted from the symbols. The coefficient B has not been included since $B = 2\bar{A}$. The numerical values printed out by the computer from the coefficients programme are given to nine decimal figures. These are, of course, not all significant and the values given Table IV have been rounded off to four or five figures which will be significant in most cases.

3) Predicted displacements.

Corresponding to each of the rays of the pencils in Table III, a set of predicted displacements $\epsilon'[2]$, $\epsilon'[3]$ were computed from the first two and first three orders of aberration coefficients respectively. As mentioned before, resulting from the choice of coordinates used for the coefficients, these displacements in the ideal image plane are for an object plane at infinity.

Diagrams

The information obtained from the computations described in the preceding paragraphs is presented in a series of diagrams. These are of three types, 1) tangential curves, 2) annular curves, and 3) spot diagrams.

1) Tangential curves

These curves show the aberrations of rays lying in the meridional plane for various values of ρ , where ρ is the radius of the zone of the aperture in which the ray lies, measured in the first polar tangent plane of the system. The vertical axis of these curves is actually marked in units of $\bar{\delta}$, but, for tangential rays this is a direct measure of ρ . The particular value of ρ corresponding to the unit $\bar{\delta}$ for each system can be ascertained from Table III.

In these diagrams three curves are drawn. The continuous line shows the variation with ρ of the true aberration ϵ'_y , determined by ray tracing. The dotted line shows the corresponding variation of the predicted displacement $\epsilon'_y[2]$, and the dashed line the predicted displacement $\epsilon'_y[3]$. The difference between the dashed and dotted line shows directly the effect of the inclusion of the seventh order coefficients in the prediction of the ray displacements.

2) Annular curves.

The annular curves are a plot of the y and z components of the displacement of rays which lie on one zone of the aperture i.e. on the rim of a circle centred on the intersection point of the principal ray with the first polar tangent plane. In each diagram, the circle chosen is that one which best fills the (equivalent) entrance pupil and which is not vignetted by more than 25 percent of its circumference. The corresponding curves for $\epsilon'_y[2]$ and $\epsilon'_y[3]$ are plotted using the same code as for the tangential curves. The radius of the circle is shown on the diagrams in units of $\bar{\delta}$.

It is convenient at times to refer to the polar coordinates of a point in the equivalent entrance pupil (i.e. in the first polar tangent plane). The point whose y, z cartesian coordinates are $+n\bar{\delta}, 0$ is designated as having polar coordinates $n\bar{\delta}, 0^\circ$. Positive rotation is taken in the direction from the positive y axis to the

positive z axis i.e. anticlockwise when looking along the direction of the positive x axis. The arrows on the annular curves indicate the direction of rotation around the curves corresponding to a positive rotation in the equivalent entrance pupil, the arrows starting at the point corresponding to 0^0 .

3) Spot diagrams.

For each pencil, the spot diagrams are plotted in pairs. Those in the column marked ε'_k are obtained from ray traces and those in column $\varepsilon'[3]$ from the first three orders of coefficients. Spot diagrams predicted by only the first two orders of coefficients have not been plotted since the appearance of these can be inferred from the behaviour of the $\varepsilon'[3]$ set in conjunction with the tangential and annular curves.

In all the diagrams, resulting from the definition of ε'_k the origin of the coordinate axes is the ideal image point appropriate to the particular pencil. The positive direction for ε'_x in the annular curves and spot diagrams is to the right of the ideal image point, and positive ε'_y is downwards. This corresponds to viewing the image plane from the lens side.

In the tangential and annular curves, where the dotted ($\varepsilon'[2]$) or dashed ($\varepsilon'[3]$) curve is not visible, they are to be considered coincident with the curve for ε'_k .

In each system, the scale of the annular curves and spot diagrams is identical with the scale of ε'_k shown on its tangential curves.

13. Quality of Agreement between ϵ_k and $\epsilon'[2]$, $\epsilon'[3]$

In this section the diagrams of each system are examined in order to determine over what range of pencil angles the predicted displacements $\epsilon'[2]$, $\epsilon'[3]$ are in good agreement with ϵ_k . In this context, the displacements $\epsilon'[2]$, $\epsilon'[3]$ computed from the coefficients will be termed secondary or tertiary predictions respectively, and the displacement ϵ_k obtained from ray traces will be termed true displacements.

Before proceeding with the examination of the diagrams, it is necessary to comment on how the comparisons are made. At first, an attempt was made to make the comparison of the various curves and spot diagrams on a quantitative basis. In particular, some of the spot diagrams were treated in the following manner. First, centroids of a set of true and tertiary predicted spot diagrams were computed. The distance of the centroids from the ideal image point gives some measure of "apparent distortion" as distinct from distortion measured from the intersection point of the principal ray which may not be in the region of the brightest part of the image. However, a single point does not enable the shape of the image patch to be determined. Accordingly, the centroid was taken as the origin of a set of rectangular cartesian axes, the y axis of which lay in the meridional plane. The z axis then splits each diagram into two parts, termed upper and lower, the upper section being part nearest the axis of the system in the case of oblique pencils, or in the direction of $-\epsilon'_y$ in the case of an axial pencil. Now, all the spot diagrams are symmetrical about the meridional plane, thus the shape of the image patch can be ascertained by considering only one of the symmetrical halves of the diagram. The coordinates of the centroids of the upper and lower "quadrants" on one side of the y axis then give some idea of the shape of the image patch, and a comparison of these coordinates between corresponding predicted and true spot diagrams gives one means of measuring the agreement between them.

This method has the advantage that all points are equally weighted. While it may have an application in some form of automatic lens design technique to detect improvements in spot diagrams as a result of changes in the system, the method proved too sensitive for the present work since it was found difficult to assign "tolerances" corresponding to "good" and "bad" agreement between the corresponding diagrams.

The decision as to whether a pair of curves or spot

diagrams are in good agreement varies according to the person making the judgment. Thus, even though a suitable formula may be developed by which to make the necessary comparisons, the tolerances to be applied would still vary from person to person. In view of this, the comparisons made here will be subjective in nature. In judging the spot diagrams, the work of Herzberger¹¹ and Y. Ukita and J. Tsujiuchi¹² will be kept in mind. In these papers are published geometrical spot diagrams and the corresponding photographic images produced by the systems. The shapes of the images are fairly well indicated by the diagrams, but in many cases the intensity distributions are not clearly predicted by the spot densities. This could be due to the techniques by which the photomicrographs were obtained, or, perhaps of more importance, due to the effect of non-linearity in the film characteristics. Also, when the dimensions of the spot diagram approach the size of the Airy disc, it might be expected that diffraction effects will seriously modify the predicted geometrical picture. Thus the spot diagrams presented here will be compared mainly on the basis of shape, the distribution of densities being a secondary consideration.

The complete set of diagrams for any one system follow the page containing the description of their behaviour.

1) Vega, fig. 10.

From an examination of the tangential and annular curves, it is obvious that the agreement between the tertiary and true curves is excellent up to the 30° pencil. At 35° the tertiary prediction is adequate, but has become poor at 40° . The secondary predictions are satisfactory up to the 30° pencil, but are starting to differ noticeably at the 35° pencil. At 40° and 45° the secondary and tertiary predictions are totally inadequate.

The true and predicted spot diagrams are in good agreement up to the 35° pencil. Then, as is so in the tangential and annular curves, the quality of the predicted spot diagrams falls off rapidly by the time the 40° pencil is reached.

The 45° annular curves have not been plotted since it is obvious that nothing can be gained from them.

Summarising, it can be stated on the basis of this system, that for systems of small aperture ($f/6.3$) the secondary predictions give a sufficiently accurate description of the aberrations up to a semi-field of 30° , and the tertiary predictions up to 35° .

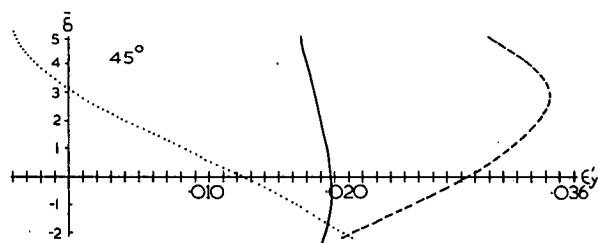
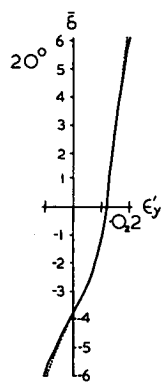
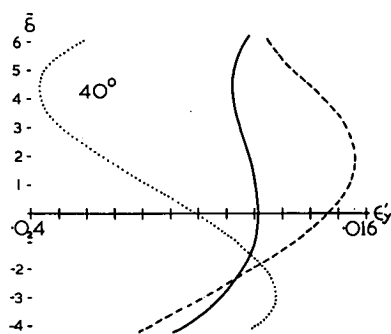
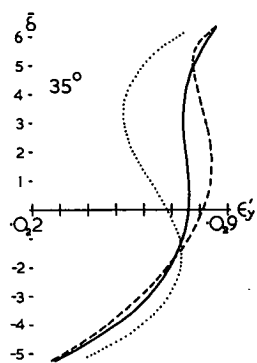
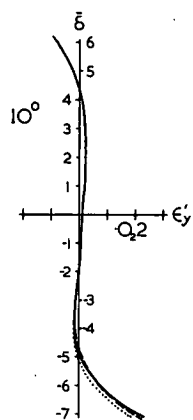
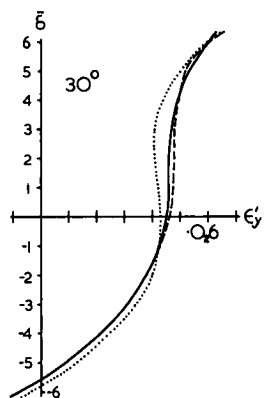
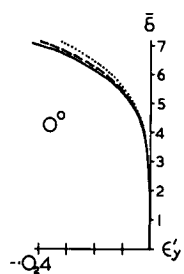


Fig. 10 Vega

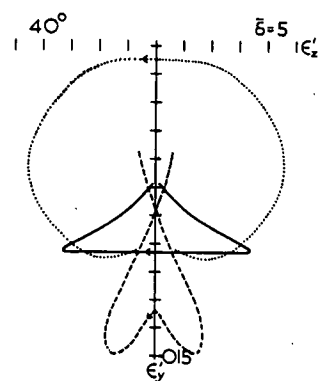
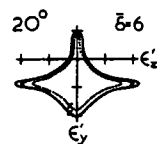
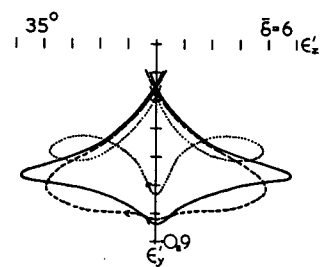
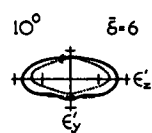
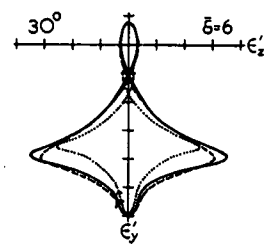
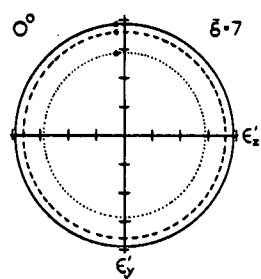
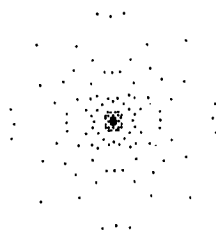
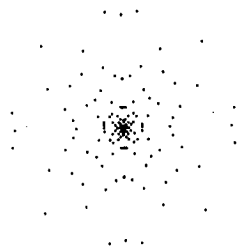


Fig. 10 Vega

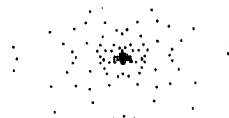
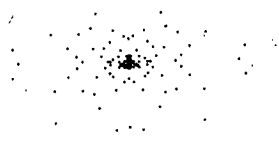
ε'_k

$\varepsilon'[3]$

0°



10°



20°



30°



Fig. 10 Vega

ϵ'_k

$\epsilon'[3]$

35°

40°

45°

Fig. 10 Vega

2) Sonnar, fig. 11

The tangential curves for this system exhibit a rather unusual feature. In the 5° and 10° pencils the tertiary curve over-shoots the true curve in the region of $\bar{\delta} < -9$. In addition in the 10° pencil overshoot occurs when $\bar{\delta} > +6$. This behaviour will be discussed later. In the 15° pencil the secondary curve is in closer agreement to the true curve than the tertiary prediction in the region of large negative $\bar{\delta}$. The curves for the 20° and 25° pencils are "normal" in the sense that the tertiary predictions are better than the secondary even though the tertiary curves themselves are not sufficiently close to the true curves.

Apart from the overshoot which occurs in extreme zone of the aperture, the tangential tertiary predictions are good up to the 15° pencil. The secondary predictions improve in quality up to the 15° pencil. At 20° and 25° , neither secondary or tertiary predictions show the shape of the true curve.

In the annular curves, the tertiary predictions are good up to the 15° pencil. As before, the secondary predictions improve in accuracy as the 15° pencil is approached.

The spot diagrams reflect the behaviour indicated in the tertiary tangential and annular predictions that is, agreement is good up to 15° , at 20° the agreement is fair and at 25° , poor.

Thus for systems of this type, average semi-field (25°) and wide aperture ($f/1.6$), the tertiary and secondary predictions are quite adequate up to 15° , and the tertiary prediction may even be satisfactory up to 20° as judged from the spot diagrams and annular curves.

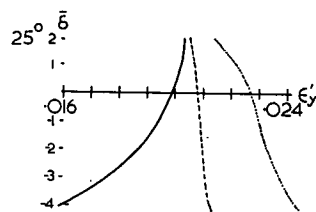
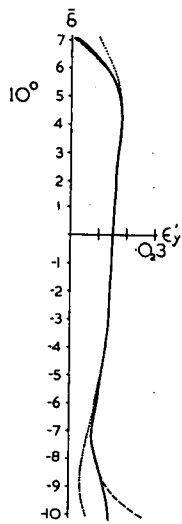
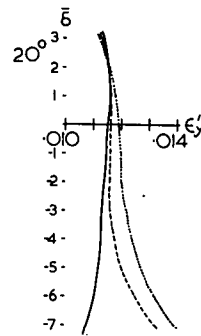
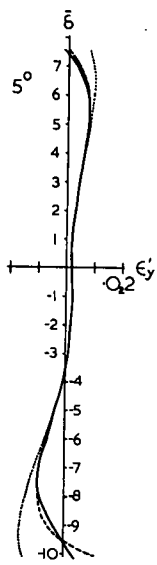
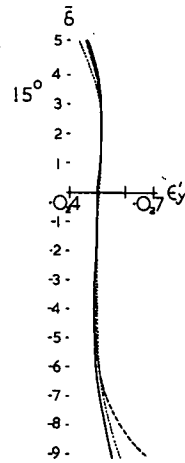
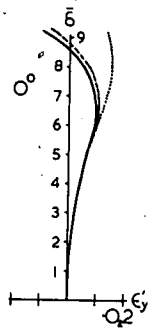


Fig. 11 Sonnar

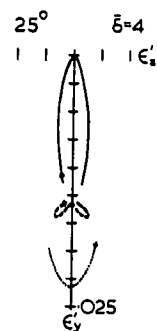
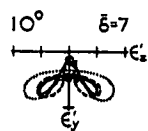
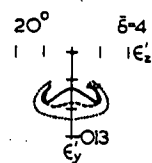
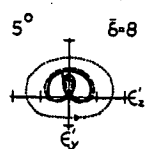
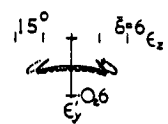
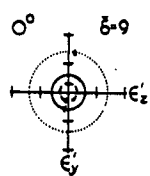


Fig. 11 Sonnar

ϵ'_k

$\epsilon'[3]$

0°



5°



10°



Fig. 11 Sonnar

ϵ'_k

$\epsilon'[3]$

15°

20°

25°

Fig. 11 Sonnar

3) Tessar, fig. 12

For this system, the tertiary tangential curves are in excellent agreement with the true curves up to the 15° pencil. The shape of the 20° and 25° true curve is adequately predicted by the tertiary curve, and is sufficiently close at 20° to be adequate for design purposes. The tangential secondary curves are adequate up to 15° , becoming bad at 25° .

The annular curves show similar behaviour. The tertiary curves maintain the shape of the true curve for all pencils, being in close agreement up to the 15° pencil. The secondary predictions are adequate up to the 10° pencil. The apparently large discrepancy between the true and secondary curve in the axial pencil is not representative of the quality of the secondary predictions for all ρ , as can be seen in the tangential curves. The secondary curves differ noticeably in shape from the 20° pencil onwards.

The spot diagrams agree very well as regards shape in all pencils. The only differences are in the spread of the points, but these are not large enough to cause concern.

Thus, for systems of this type, the tertiary predictions are sufficiently accurate up to at least 25° semi-field, and the secondary predictions up to about 15° .

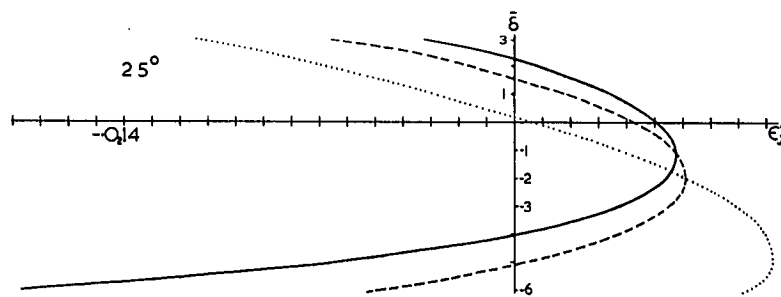
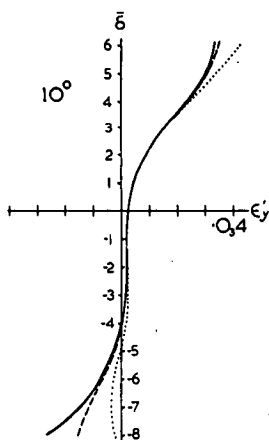
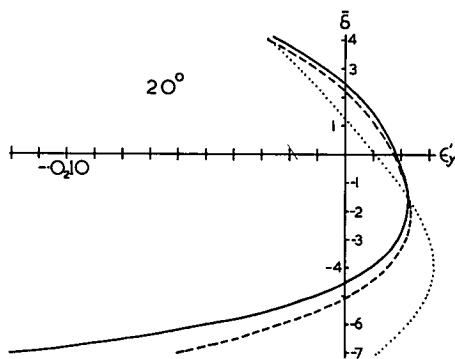
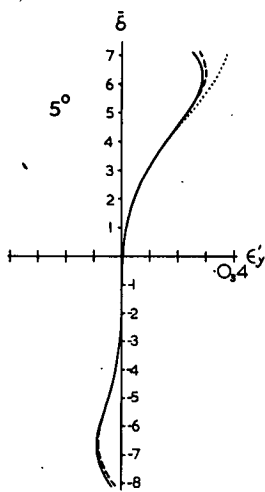
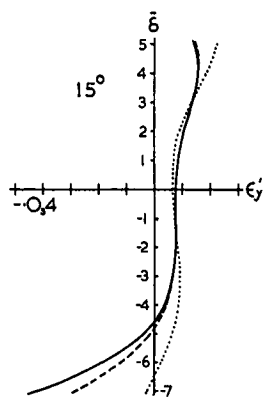
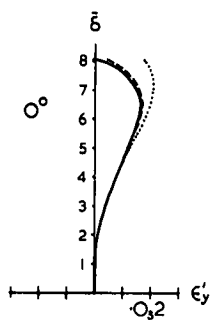


Fig. 12 Tassar

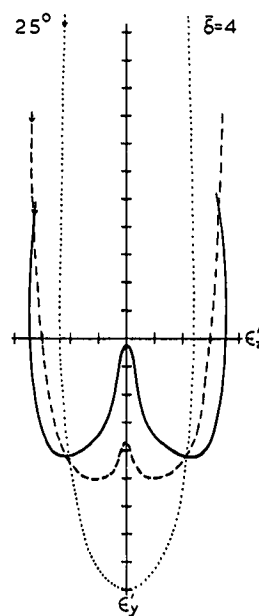
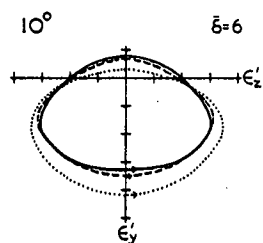
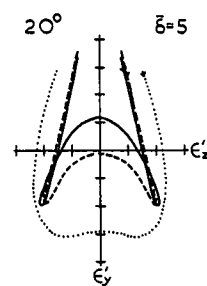
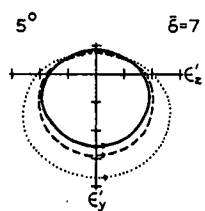
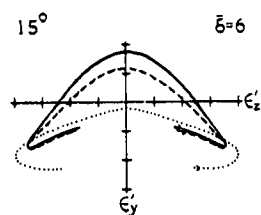
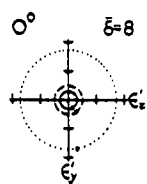


Fig. 12 Tessar

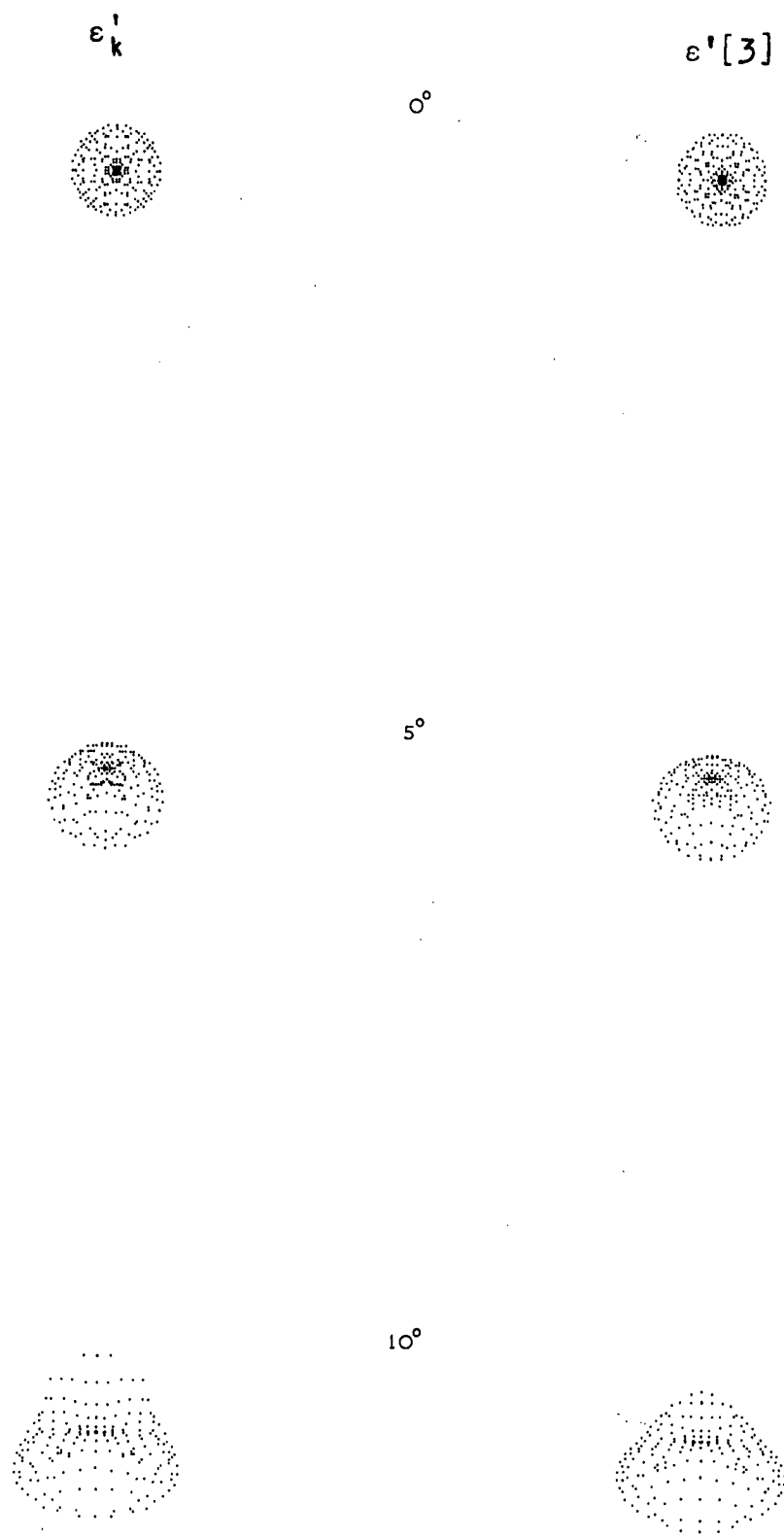
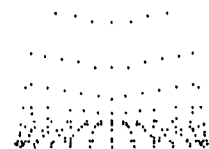
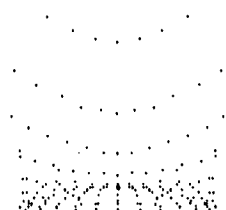


Fig. 12 Tessar

ϵ'_k

$\epsilon'[3]$

15°



20°



25°

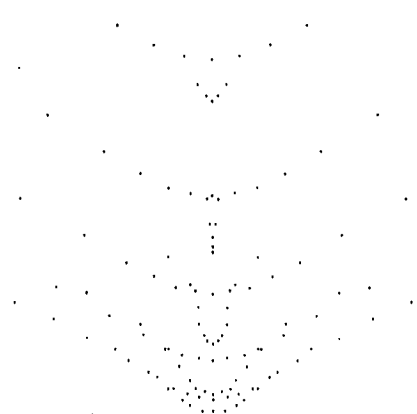
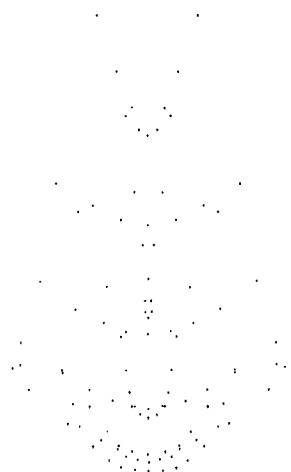


Fig. 12 Tessar

4) Petzval, fig. 13

In all the diagrams, the tertiary predictions are in excellent agreement with the true results for all pencils. The secondary tangential curves are close enough to the true curves to be adequate for design purposes.

In the annular curves, the secondary agreement does not appear to be good, but this is only the result of radii chosen for these curves.

For this type of system then, primary and secondary coefficients describe the aberrations with sufficient accuracy up to a semi-field of 10° . The inclusion of the tertiary coefficients effects an improvement in the quality of the predicted displacements, but are not absolutely necessary.

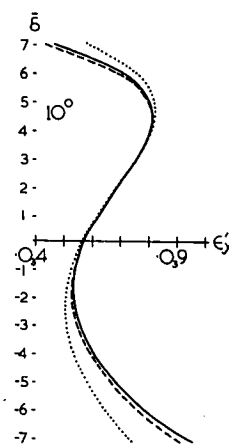
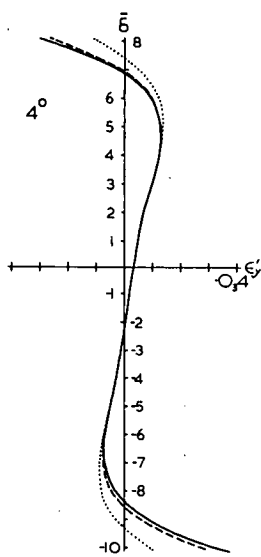
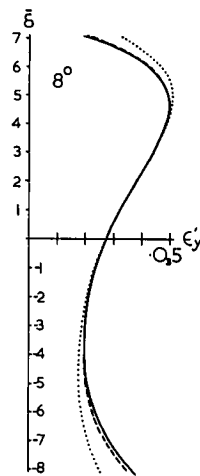
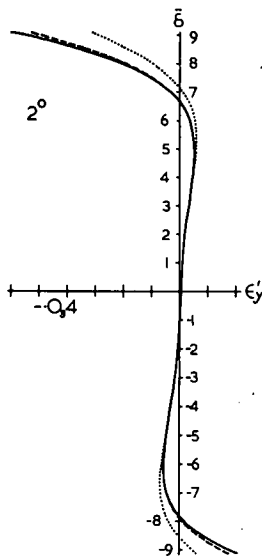
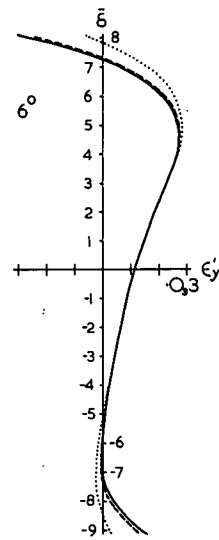
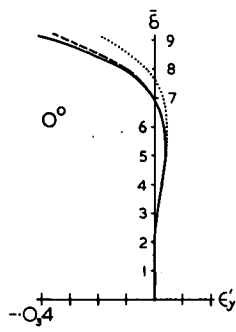


Fig. 13 Petzval

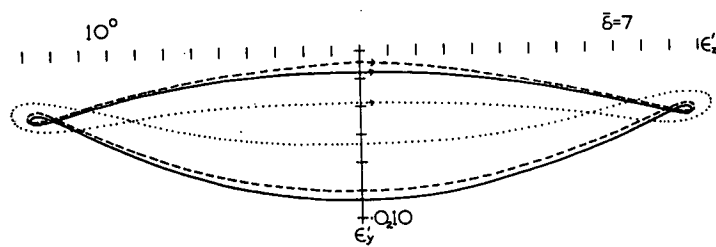
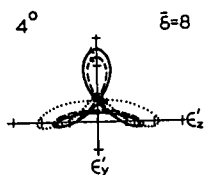
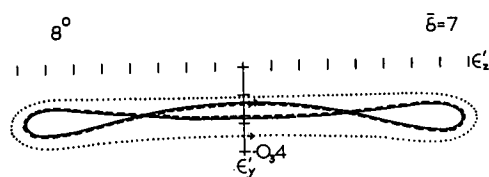
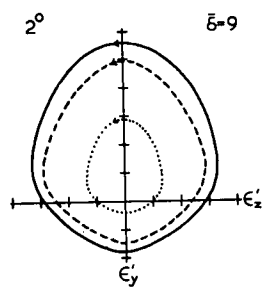
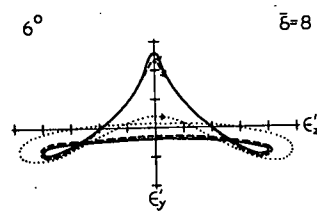
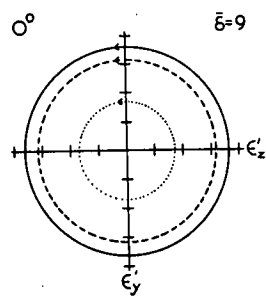


Fig. 13 Petzval

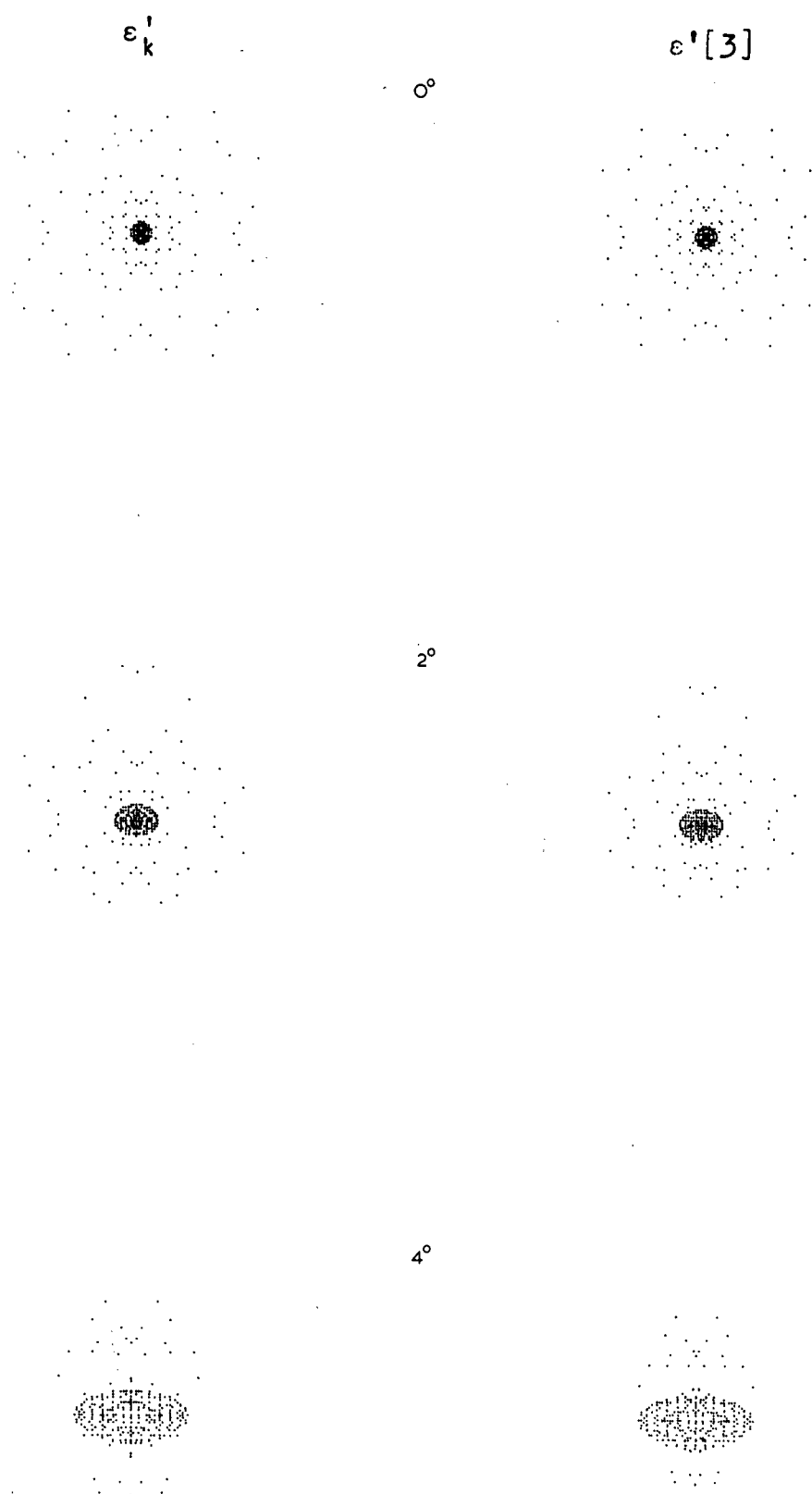


Fig. 13 Petzval

5) Telephoto, fig. 14

The tangential tertiary curves agree well over all pencils and the secondary curves have a satisfactory agreement up to the 5° pencil.

The annular tertiary predictions are quite good over all the pencils, and the secondary predictions are satisfactory up to the 6° pencil.

Consistent with the tangential and annular tertiary curves, the spot diagrams are in good agreement for all pencil angles.

Owing to the high degree of correction of this system, these diagrams are plotted on a scale larger than for the other systems. In the tangential curves, the maximum difference between the secondary and the true curve is $0.000043 \cdot f'$ in the 7° pencil. It may be argued that this difference is so small as to be insignificant and hence that the secondary curve for this pencil angle is in good agreement with the true curve. The view taken here is that since all the aberrations except distortion are small, a high degree of precision is also required by the predictions if the coefficients are to be of any use in obtaining a well corrected system.

This system was designed originally to work at a focal of 48 inches. Under test, it was noted that the performance was slightly inferior at the $4^\circ - 5^\circ$ zone compared with the centre and edge. This is in accordance with the appearance of the spot diagrams, and indicates that the system could be further improved (i.e. diffraction limit has not been reached). To have any reasonable chance of achieving this, the coefficients must very accurately represent the state of correction of the system. It will be noted that in the 7° tangential pencil the $\epsilon'[2]$ curve indicates a range of ϵ'_y twice as great as that indicated by the true curve, thus at this inclination the secondary coefficients are not sufficient. In order to "remove" the aberrations by balancing between low and high orders, the tertiary coefficients must be included to give reliable results.

ϵ'_k

$\epsilon'[3]$

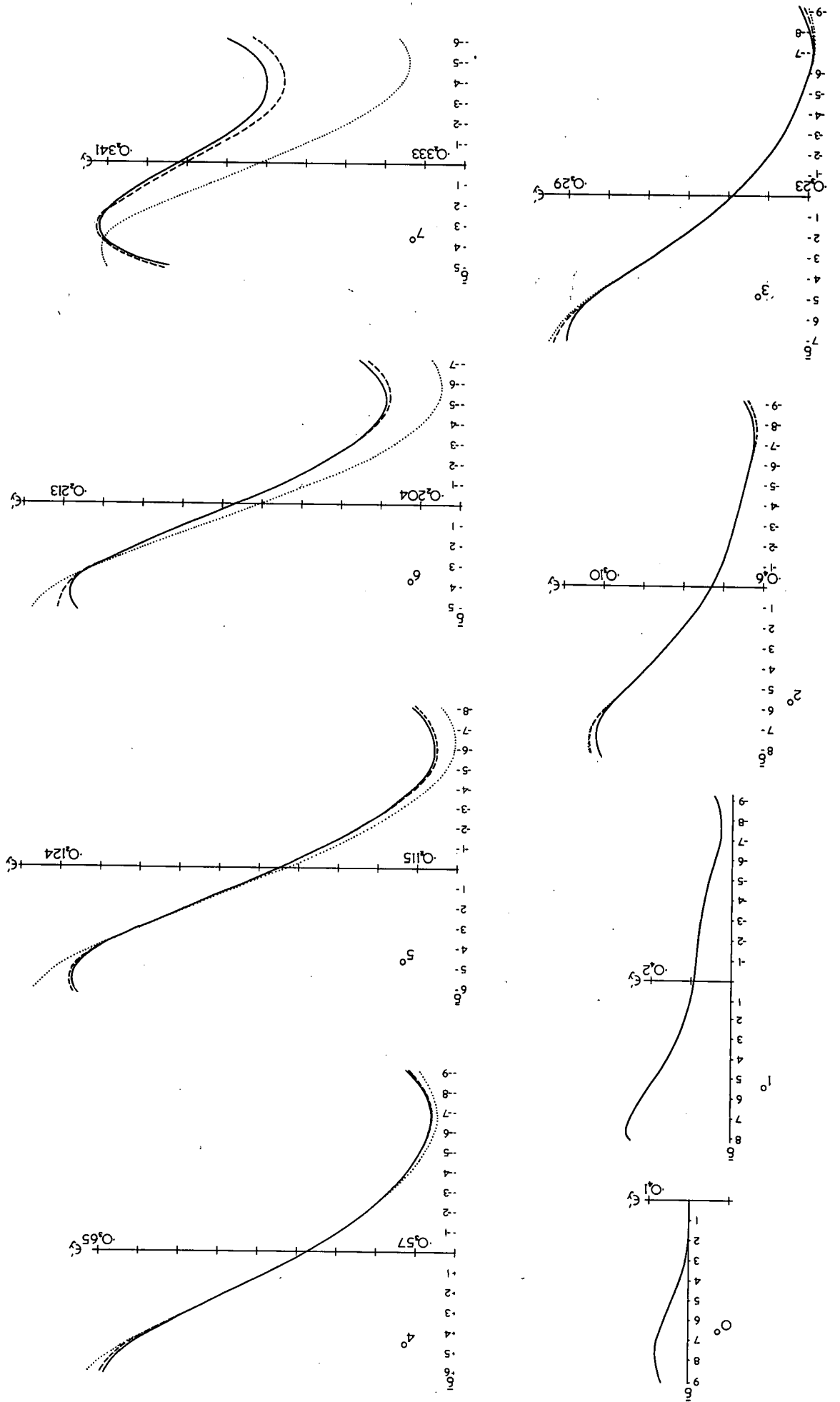
6°

8°

10°

Fig. 13 Petzval

Fig. 14 Telephoto



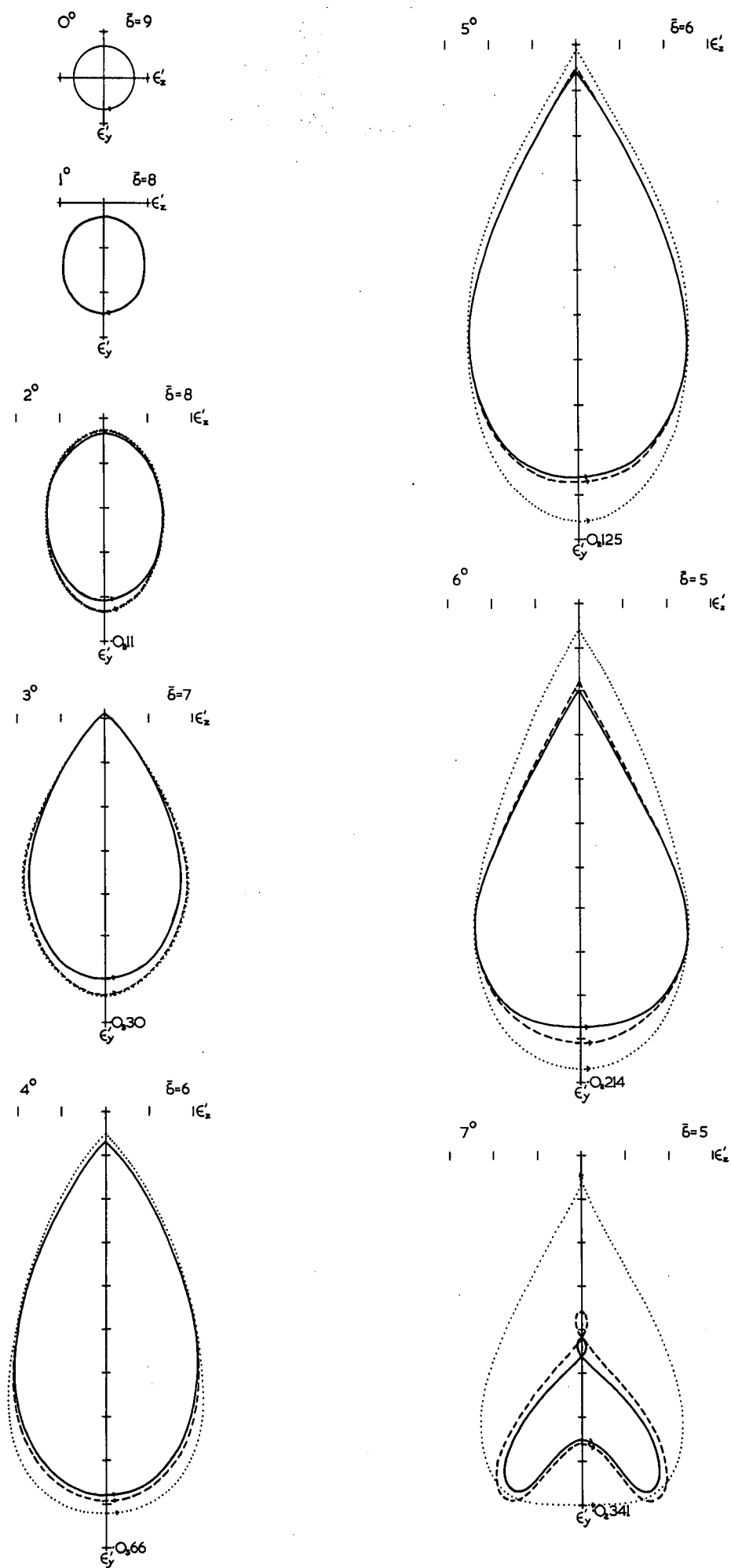


Fig. 14 Telephoto

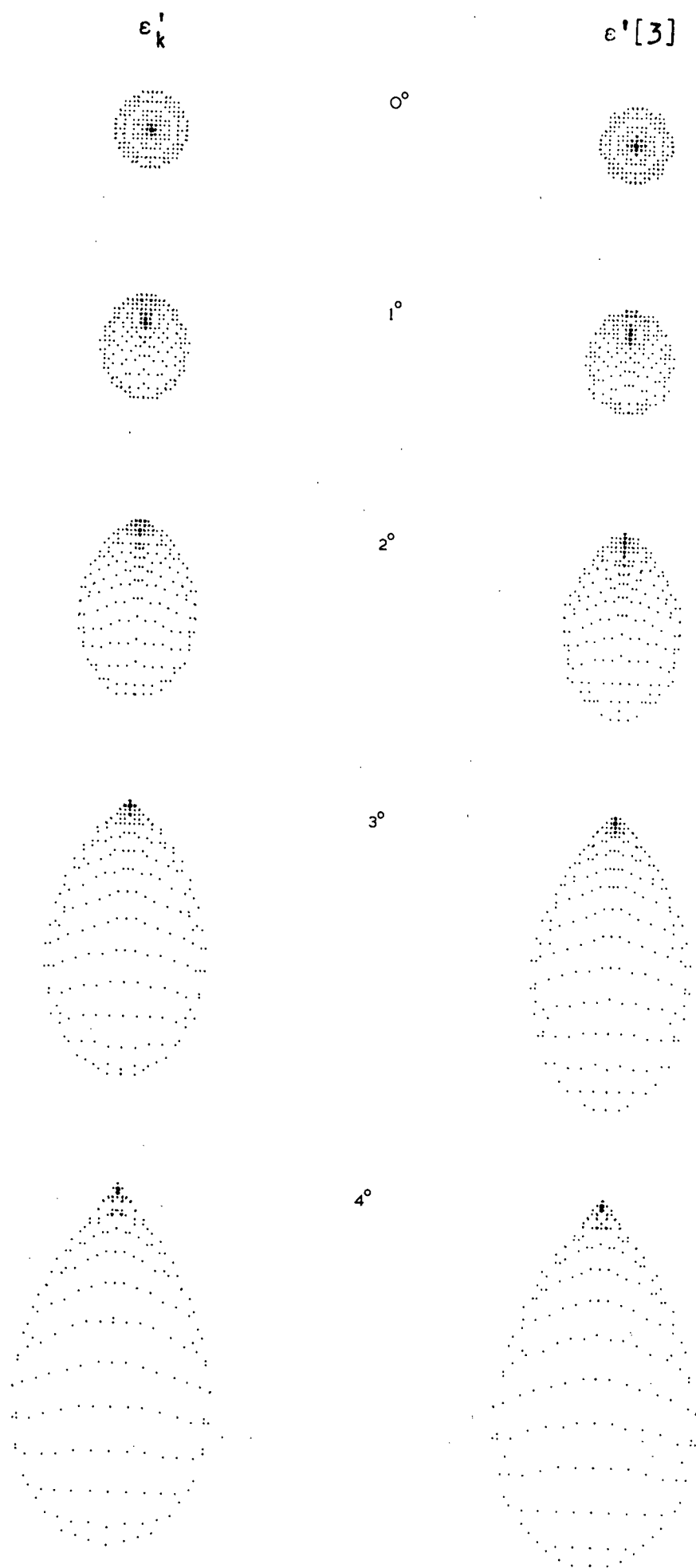


Fig. 14 Telephoto

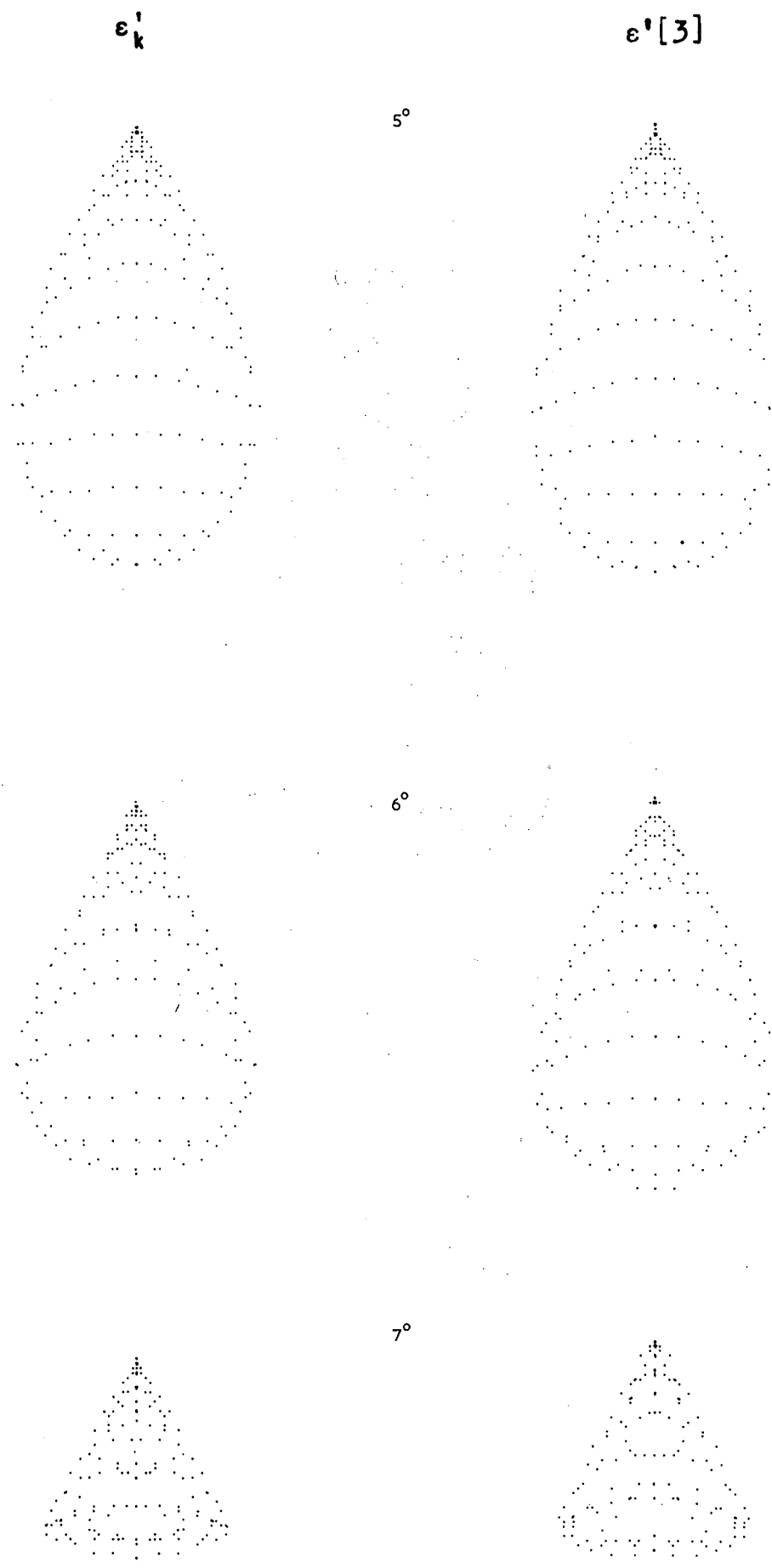


Fig. 14 Telephoto

6) Biotar, fig. 15

Over all pencils, the tangential tertiary curves are in good agreement with the true curves, except for the marginal rays in the early pencils. Not until the 9° pencil do the marginal tertiary predictions approach the true curve. The secondary tangential predictions show similar behaviour, but for these, close agreement is attained from the 15° pencil onwards.

The inaccuracy of the secondary and tertiary predictions for marginal rays is very clearly indicated by the 3° annular curves. At 9° the tertiary predictions are considerably better, and from this pencil onwards, the tertiary curves show satisfactory agreement. Only in the 18° pencil does the secondary curve give any indication of the true shape.

On the other hand, the spot diagrams do not differ significantly from one another in any pencil. This, once again, indicates that inaccurate predictions for marginal rays are not serious, and that annular curves constructed with extreme values of $\bar{\delta}(\rho)$ can lead to erroneous conclusions.

Thus, tertiary coefficients are needed to describe adequately the aberrations of this type of system up to 18° semi-field. Primary and secondary coefficients alone will only give sufficient indication of the behaviour of the system between 12° and 18° semi-field.

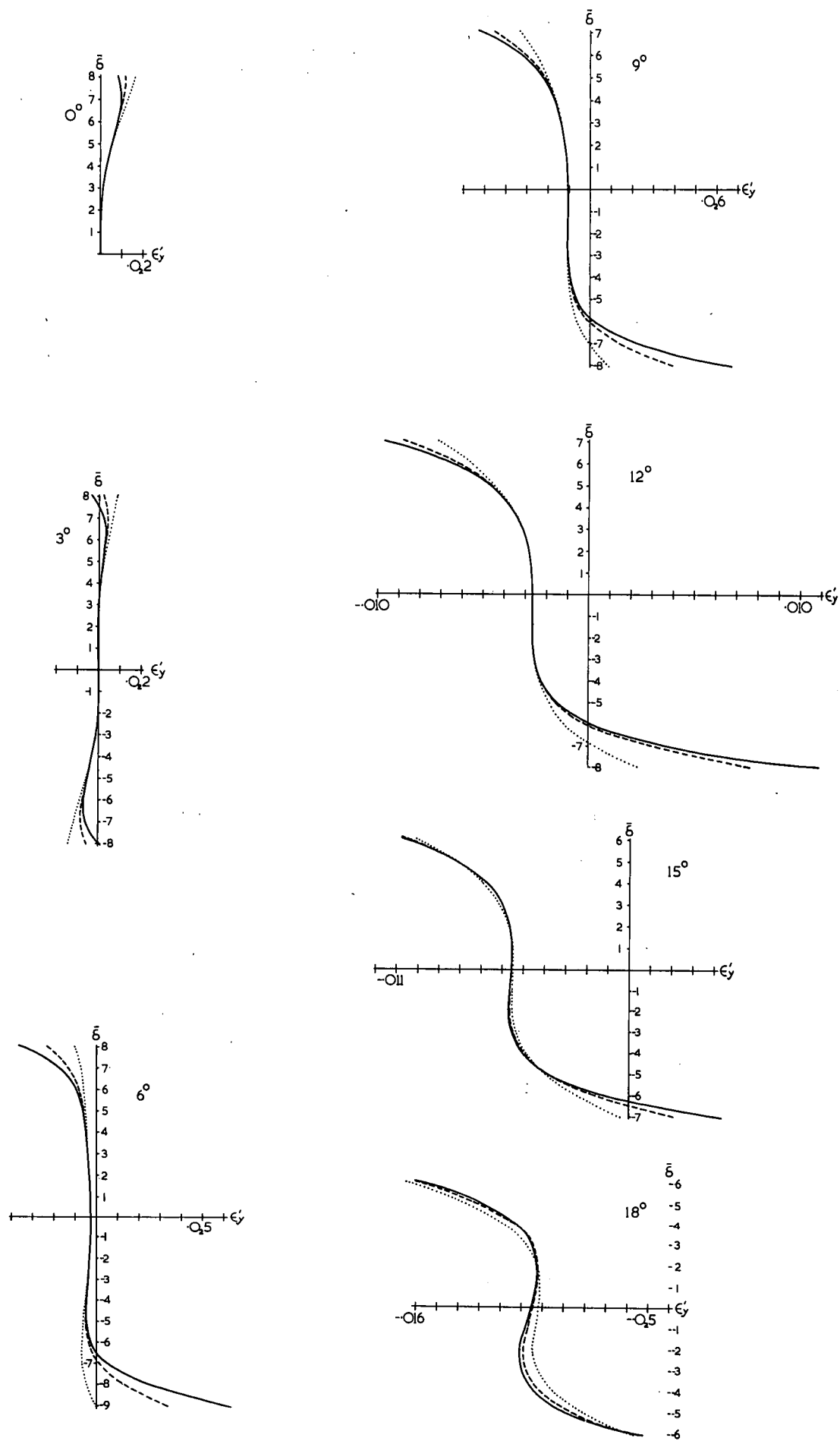


Fig. 15 Biotar

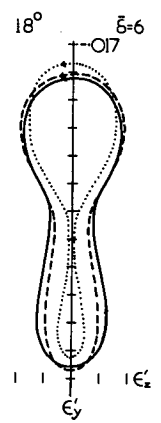
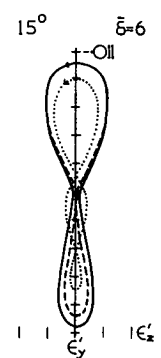
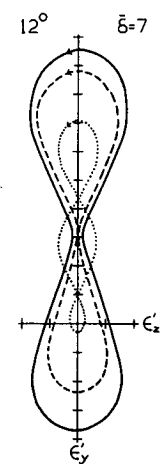
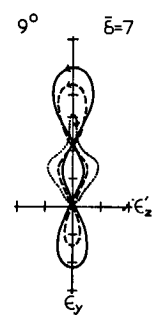
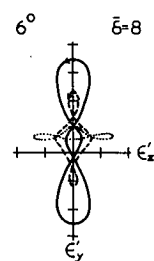
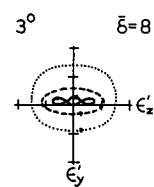
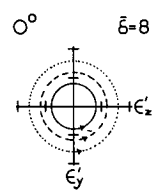


Fig. 15 Biotar

ε'_k

$\varepsilon'[3]$

0°



3°



6°



9°



Fig. 15 Biotar

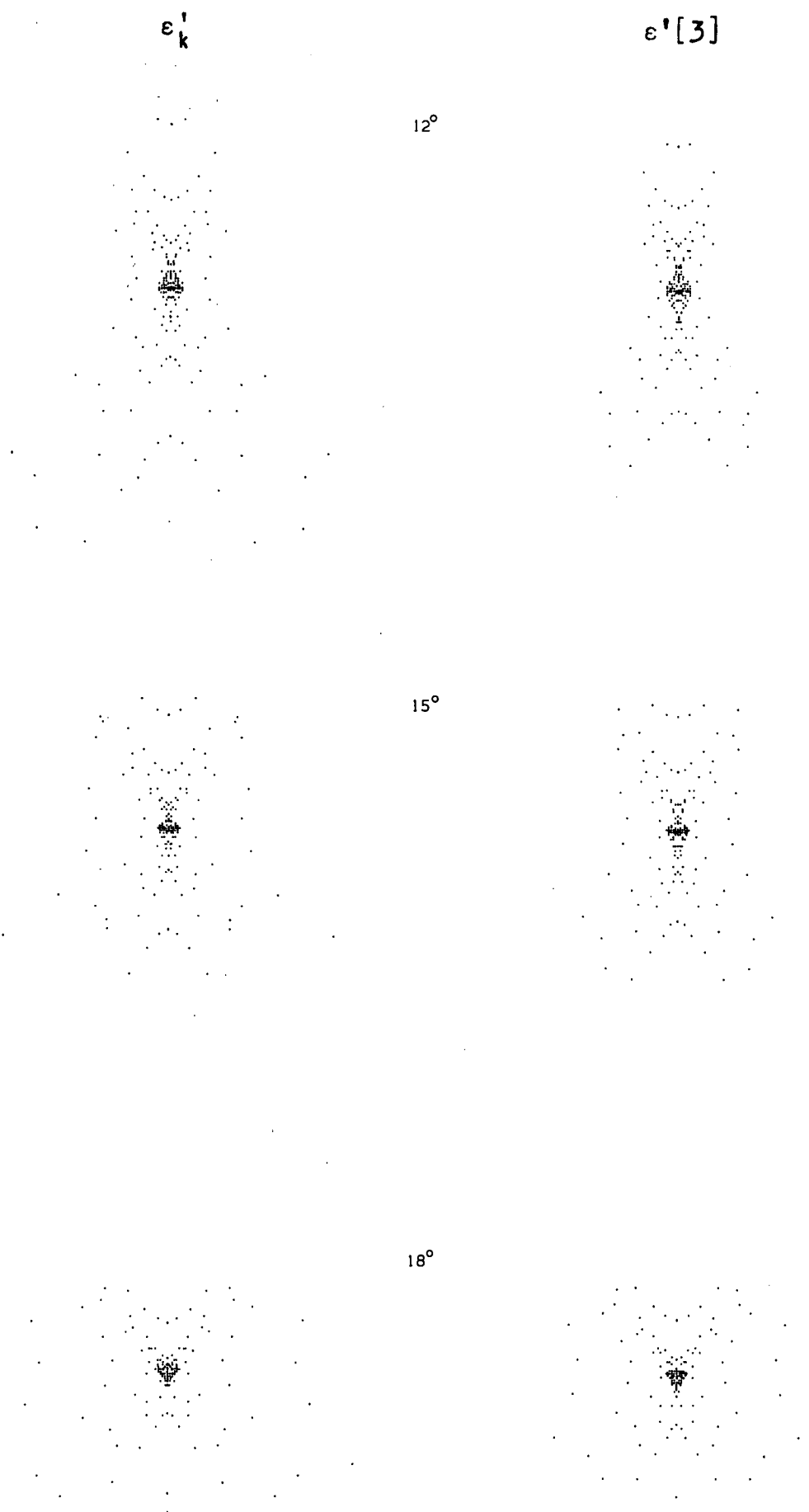


Fig. 15 **Biotar**

Some further comments on the behaviour of the curves for the Sonnar and Biotar are relevant at this point.

In both of these systems, the tertiary tangential curves differ markedly from the true curves when $\bar{\delta}$ becomes large in pencils for the first half of the field. It will be also noticed that both of these systems have maximum stop numbers $f/1.6$, being the largest apertures of all the systems considered here. From an examination of the Biotar tangential curves, it is apparent that part of the discrepancies between the tertiary and true curves can be accounted for by higher order spherical aberration (ninth order and higher). Using the 0° pencil to obtain an estimate of the effect of the higher orders for a $\bar{\delta} = 8$, the 3° annular curves were redrawn, the predicted curve now including the effect of all orders of spherical aberration (fig. 16). It will be seen that the agreement is now considerably better than before. From an examination of the 3° tangential curves, allowing for all orders of spherical aberration, it appears very likely that the remaining predominant aberration necessary to account for the discrepancies will be ninth and higher orders of oblique spherical aberration, with small amounts of higher order circular coma. The reason for choosing these aberrations is based on the fact that the discrepancies are being caused by higher order aberrations which have a comparatively large effect for large values of ρ ($\bar{\delta}$) and for small values of pencil angle (\bar{H}). Apart from spherical aberration, which does not account for all of the discrepancies, n -th order circular coma and oblique spherical aberrations, the coefficients for which multiply $\rho^{n-1} \bar{H}$ and $\rho^{n-2} \bar{H}^2$ respectively, will be the next most predominant aberrations in the region of interest. Higher order circular coma is not likely to be large since the asymmetry of the true 3° annular curve is not great. While it is possible to obtain an estimate of the effect of higher order spherical aberration from the axial tangential curves, it is not possible to estimate easily the effect of other higher order aberrations (except distortion). If one particular higher order aberration is assumed predominant, then its effect can be estimated, but the usefulness of this estimate is dependent on the correctness of the assumption.

The annular curves for $\bar{\delta} = 7$ corresponding to an aperture of $f/1.8$ show that the tertiary curve is in reasonable agreement with the true curve (fig. 17) indicating that the unknown residual aberrations have a marked effect only at the maximum ρ , supporting the reasoning already presented.

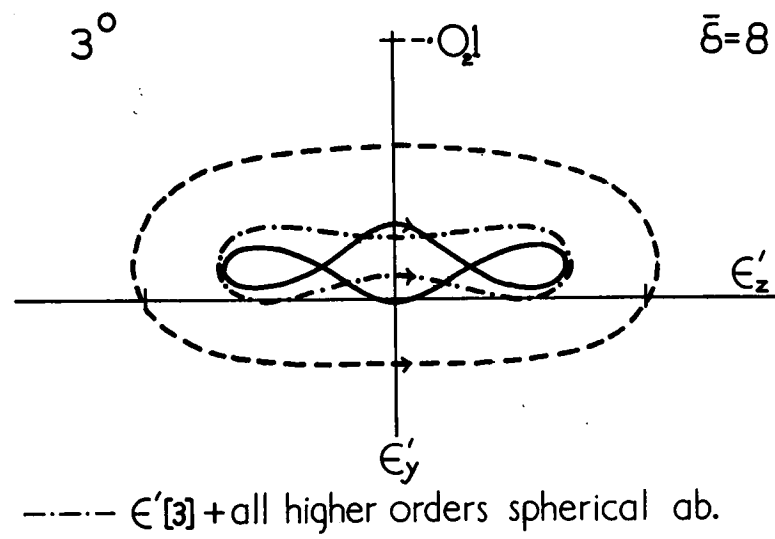


Fig. 16 Biotar 3^0 annular curve

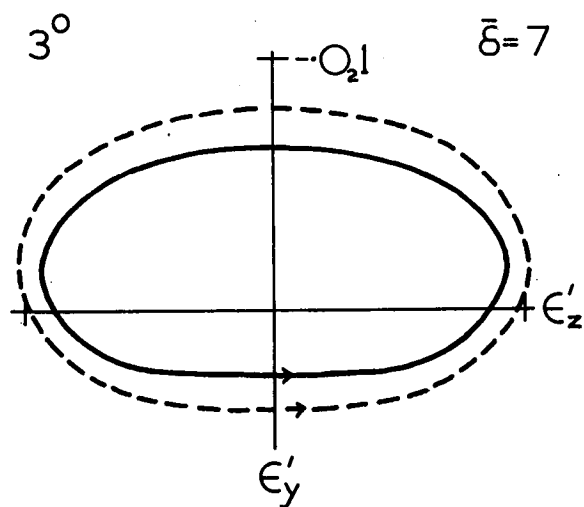


Fig. 17 Biotar 3^0 annular curve with $\bar{\delta}=7$

In the Sonnar 5° and 10° tangential curves, the tertiary curve crosses the true curve at $\bar{\delta} = -10$. In the 15° pencil, the addition of the tertiary coefficients causes the predicted values $\epsilon'[3]$ to become worse than $\epsilon'[2]$ in the region of $\bar{\delta} < -6$. As in the case of the Biotar, these effects are due to the omission of the 9th and higher order aberrations. The axial tangential curves indicate that the predominant higher order spherical aberration coefficients are negative. The 9th and 11th order spherical aberration coefficients for this system have been computed and they are -30.169 and -53.71 respectively. However, the addition of these, while improving the agreement in the region of positive $\bar{\delta}$, would make the agreement worse at the point $\bar{\delta} = -10$ in the 5° and 10° pencils. This system, on further examination, exhibits a peculiar behaviour. When the value of $\bar{\delta}$ is extended beyond 9 in the axial pencil, the true curve very rapidly swings around and recrosses the axis, ϵ'_y becoming very large and positive in the space of a small increase of $\bar{\delta}$. This behaviour is just starting to appear in the 5° and 10° pencil at $\bar{\delta} = -10$.

By using eq. (7.2) in O.A.C. VIII, reference no. 9 to compute the effect of higher orders of spherical aberration at those surfaces whose contributions to this aberration are dominant, the true axial curve can be satisfactorily predicted for values of $\bar{\delta}$ in excess of 9. It appears that quite high orders may have to be reached (perhaps 21^{st}) before the coefficients become sufficiently positive to make the predicted curve follow the true curve in the region of large ρ . However, in the region of interest here, $f/1.6$, the spot diagrams are in quite good agreement. It will be noticed that the annular curves are plotted for a maximum ρ corresponding to $f/1.9$, thus avoiding the region of large deviations of the tertiary tangential curves from the true ones.

Thus, although the higher order coefficients of spherical aberration have a marked effect at large values of $|\bar{\delta}|$, the tertiary predictions are quite adequate for pencils up to 15° .

14. Criteria Indicating Accuracy of Predicted Displacements.

The purpose of this section is to discover what simple criteria, if any, will indicate the quality of the agreement between predicted and true displacements. Since we are attempting to estimate ϵ_k^i from only the first two or three orders of its infinite series expansion (3.17), the accuracy of the estimate will depend on the rate of convergence of the series. So far it has proved difficult to determine theoretically the rate of convergence, so that conditions for good agreement by the secondary and tertiary predictions are here and elsewhere obtained by empirical methods. These methods involve an examination of the various angles which arise during the tracing of the paths of selected tangential rays for several pencil inclinations, keeping in mind the quality of the predictions for each pencil.

If the algebraic ray trace scheme described in Section 9 is carried out with a desk machine, the information necessary to compute the various angles associated with the ray at each surface is readily available. In particular, for tangential rays, at any surface

$$V = \tan U_y \quad (14.1)$$

$$\underline{P} = \cos (I + U_y) \quad (14.2)$$

where U_y is the angle of inclination of the ray with the axis, and I is the angle of incidence at the surface. (I here is the Conrady symbol, and should not be confused with the $I (=cY - V)$ of the algebraic theory (eq. 2.1)). The programmed ray trace does not make available any information relating to the individual surfaces so that it was necessary to obtain the required angles by hand ray tracing. For this purpose, it was found that a trigonometric ray trace was somewhat faster than the algebraic method and had the added advantage that the angles required appeared directly in the course of computation. The gain in speed occurs only when tangential rays are traced. In fact, the particular trigonometric method used will only trace tangential rays whereas the algebraic method is faster for a computer and works equally well for skew rays.

To assist in the examination of the results a table indicating the quality of the agreement between the various predictions and true results for all the pencils of all the systems has been compiled (table V) based on the comments of Section 13. In this table, for each system, the inclination of the pencils in degrees is given in the row marked V_1 . The agreement of the secondary curves is

TABLE V

Quality of agreement between diagrams

Vega	V_1^O	O	10	20	30	35	40	45	
	T_2	G	G	G	F	F	B	B	
	T_3	G	G	G	G	F	B	B	
	A_2	G	G	G	F	F	B	B	
	A_3	G	G	G	G	F	B	B	
	S_3	G	G	G	G	G	F	B	
Sonnar	V_1^O	O	5	10	15	20	25		
	T_2	F	F	F	G	B	B		
	T_3	G	G	G	F	B	B		
	A_2	B	B	F	F	B	B		
	A_3	G	G	G	G	F	B		
	S_3	G	G	G	G	F	B		
Tessar	V_1^O	O	5	10	15	20	25		
	T_2	F	F	F	F	B	B		
	T_3	G	G	G	G	F	F		
	A_2	B	F	F	B	B	B		
	A_3	G	G	G	G	F	F		
	S_3	G	G	G	G	G	G		
Petzval	V_1^O	O	2	4	6	8	10		
	T_2	F	F	F	F	F	F		
	T_3	G	G	G	G	G	G		
	A_2	B	B	B	B	B	B		
	A_3	G	G	G	G	G	G		
	S_3	G	G	G	G	G	G		
Telephoto	V_1^O	O	1	2	3	4	5	6	7
	T_2	G	G	G	G	G	G	F	F
	T_3	G	G	G	G	G	G	G	G
	A_2	G	G	G	G	G	G	F	B
	A_3	G	G	G	G	G	G	G	G
	S_3	G	G	G	G	G	G	G	G
Biotar	V_1^O	O	3	6	9	12	15	18	
	T_2	B	B	B	F	F	F	F	
	T_3	F	F	F	G	G	G	G	
	A_2	B	B	B	B	B	B	F	
	A_3	F	B	F	G	G	G	G	
	S_3	G	G	G	G	G	G	G	

TABLE VI

Vega					Sonnar				
V_1^0	$\bar{\delta}$	I	I'	U	V_1^0	$\bar{\delta}$	I	I'	U
0°	7	38 _s	37 ₄	20 ₄	0°	9	50 ₉	52 ₉	31 _s
	5	26 _s	13 ₄	13 ₄		7	36 ₉	37 ₉	23 _s
10°	7	39 _s	34 ₄	32 ₄	5°	7	35 ₉	36 ₉	31 _s
	5	27 _s	23 ₄	26 ₄		0	9 ₇	8 _s	9 ₇
	0	8 ₇	9 ₂	14 ₇	10°	-10	57 ₉	61 ₉	26 _s
	-5	24 _s	26 ₄	22 ₆		6	38 _s	32 _s	36 _s
	-7	35 _s	38 ₄	26 ₆		0	18 ₇	15 _s	18 ₇
20°	7	38 _s	33 ₆	37 ₃	15°	-9	50 ₉	52 ₉	13 _s
	0	17 ₇	17 ₂	29 ₇		5	41 _s	35 _s	39 _s
	-7	32 ₃	38 ₄	28 ₆		0	28 ₇	22 _s	27 ₇
30°	7	38 ₇	35 ₆	54 ₄	20°	-9	53 ₉	55 ₉	23 ₇
	0	24 ₇	25 ₂	43 ₇		3	44 _s	37 _s	42 _s
	-7	35 ₃	37 ₄	50 ₆		0	37 ₇	30 ₂	37 ₇
35°	6	39 ₇	31 ₆	56 ₄	25°	-7	46 ₉	48 ₉	34 ₇
	4	35 ₇	25 ₆	52 ₄		2	49 _s	41 ₂	48 ₇
	0	27 ₇	29 ₂	50 ₇		0	48 ₇	38 ₂	47 ₇
	-4	28 ₃	36 ₂	52 ₆		-4	47 ₇	42 ₆	46 ₇
	-6	34 ₃	39 ₂	55 ₆					
40°	6	41 ₇	30 ₆	61 ₃					
	4	38 ₇	26 ₂	60 ₃					
	0	30 ₇	33 ₂	57 ₇					
	-4	-24 ₂	40 ₂	59 ₇					
	-6	36 ₃	43 ₂	61 ₆					
45°	6	40 ₇	29 ₂	67 ₃					
	3.5	38 ₇	31 ₂	66 ₃					
	0	31 ₇	37 ₂	65 ₇					
	-3.5	26 ₇	43 ₂	66 ₇					
	-6	35 ₃	45 ₂	67 ₇					

TABLE VI continued

Tessar					Petzval				
V_1^0	$\bar{\delta}$	I	I'	U	V_1^0	$\bar{\delta}$	I	I'	U
5°	7	26_3	23_7	17_3	6°	8	40_3	39_2	24_3
	0	7_3	7_2	7_3		0	12_7	13_6	8_7
15°	6	39_3	29_2	28_3		-9	22_3	28_6	9_3
	0	20_3	19_2	19_3	10°	7	43_3	42_2	28_3
	-6	35_6	33_6	20_6		0	21_7	22_6	14_7
20	4	40_3	33_2	33_3		-7	28_7	34_6	13_7
20°	0	29_3	27_2	27_3	Biotar				
	-7	42_6	40_4	28_5					
25°	3	45_3	39_2	39_3					
	0	36_3	34_4	34_3	0°	7.5	37_6	36_5	22_6
	-6	45_6	43_4	35_5		6	29_6	28_5	17_5
Telephoto					6°	8	48_6	33_8	29_5
V_1^0	$\bar{\delta}$	I	I'	U		6	36_6	24_8	23_5
5°	8	25_3	21_2	16_3		0	10_9	10_2	9_6
	6	21_3	18_2	14_3		-6	32_7	35_5	12_8
	0	13_9	17_8	13_9		-8	43_7	47_5	18_5
	-6	14_9	21_8	13_9	12°	7	49_6	33_4	32_5
-8	14_8	23_8	13_9	4		32_6	24_4	24_5	
7°	5	24_3	21_2	18_9		0	20_9	18_7	17_6
	0	19_3	23_8	18_9		-4	33_7	33_7	16_8
	-5	19_9	27_8	18_9	-7	47_7	48_7	20_8	
					18°	5	42_6	34_4	33_6
						0	33_9	30_7	28_6
						-6	55_7	56_7	29_8

presented in row T_2 , and the tertiary agreement in row T_3 . Similarly, the agreement between the annular curves is given in rows A_2 , A_3 . Finally, the comparison of the tertiary spot diagrams is given in row S_3 .

The quality of agreement is indicated by one of three letters G, F, B. G(good) is used when the shape and position of the predicted figure agrees well with the true result. F(fair) denotes that the predictions indicate approximately the shape and position of the true figure. B(bad) denotes a bad prediction of either shape or position.

In Table VI are tabulated in degrees the maximum angles of incidence, I , of refraction, I' , and inclination with the axis U , of tangential rays whose ρ is given by the value of $\bar{\delta}$ in conjunction with table III. The signs of the angles have been disregarded.

Not all the pencils corresponding to the curves and spot diagrams have been tabulated as in some systems sufficient information is given by fewer pencils. The tangential rays traced have been chosen to cover the main points of interest in the tangential curves. The subscripts after the figures indicate the surfaces at which the maximum angles occur.

The values of I for the Sonnar in Table V exceed 45° in pencils 0° , 5° , 10° , 15° , for those values of ρ where, in the tangential curves, the tertiary agreement is poor. Similarly, in the 25° pencil for the Tessar, I reaches 45° for the extreme rays of the pencil, also corresponding to the points of maximum deviation in the predicted tangential curves. At first sight, then, it would appear that the quality of the tertiary predictions is closely correlated with the value of I , or I' , (as seen in the 6° pencil at $\bar{\delta} = -8$ for the Biotar). However, in the 18° pencil for the Biotar, at a $\bar{\delta}$ of -6 , I and I' are well in excess of 45° but the tertiary agreement is quite good in this tangential pencil.

In the Vega, neither I and I' exceeds 45° for any pencil, yet the spot diagrams are in poor agreement from the 40° pencil onwards. However, U steadily increases as the pencil angle increases, tending to suggest that the magnitude of this angle may also control the quality of agreement.

In the course of the expansion of $\Delta\Delta$ as a power series in \underline{Y} , \underline{V} , quantities of the type $(1 + \underline{V}^2)^{\frac{1}{2}}$ and $(1 - \sin^2 I)^{\frac{1}{2}}$, $(1 - \sin^2 I')^{\frac{1}{2}}$ are in effect expanded as power series and the convergence of these series depends on the magnitude of \underline{V} , $\sin I$, $\sin I'$. Thus it may be

expected that quality of agreement between the predicted and actual displacements would depend to some extent on the values of I , I' , U ($\tan U = V$). However, it is the products of the above series which occur in the expansion for $\Delta\Delta$ and thus it is not likely that a simple and straight forward relation between the magnitudes of the angles and the quality of agreement can be obtained.

Nevertheless, in an attempt to find a single criterion, the ratio $\sin I / \cos U$ was examined. This ratio is equal to $I(cY-V)$, for tangential pencils, a factor which occurs in the expansion of $\Delta\Delta$ (M 60.3). In each of the systems examined here it can be shown that providing $\sin I / \cos U$ is less than 1 at each surface for the tangential rays given in table VI then the tertiary predictions are in sufficiently good agreement with the true results. (The maximum values of $\sin I / \cos U$ cannot be obtained from table VI since the angles I , I' , U are not necessarily listed for the same surface). In addition, the two systems discussed in O.A.C.V.⁶ were also examined, and here also the behaviour of the ratio was found to be in accordance with the above statement. The converse is not necessarily true, that is, $\sin I / \cos U$ can exceed 1 at some surface and yet the agreement may still be satisfactory. System Z_2 in O.A.C.V. is such a case.

While this ratio shows promise as a criterion for good agreement, it has not been possible so far to find any reasonable theoretical reason why it should be any more reliable than the value of I , I' or U . It is very likely that this ratio is just a fortuitous arrangement of angles which happens to incorporate the features already noted about the magnitudes of these angles.

To summarize, the following statements can be made. The agreement between the tertiary predicted and true displacements will be good providing that nowhere in the system does I , I' exceed 45° or U exceed 50° . It appears (with no theoretical justification) that, providing $\sin I / \cos U$ does not exceed unity anywhere in the system, the tertiary predictions will be good. Many more systems will have to be examined before the last statement can be relied upon. The converses of these statements are not necessarily true. It has been found on several occasions that the above limits can be exceeded and yet the agreement be still good.

It must be realised that these statements are based on the results of observation, in some cases coupled with slight theoretical justification. Thus they may be found to be wrong at any time.

They are intended only as a guide in the use of the aberration coefficients, giving the designer a quick means of assessing whether the first three orders of coefficients alone are likely to represent, with sufficient accuracy, the state of correction of the system.

There is, however, a more reliable way of determining the quality of agreement, which does not involve much more work. From an examination of the diagrams of the various systems, it will be noticed that when the predicted tertiary annular curves approximate the true curves, the spot diagrams are in good agreement. This is so providing that the ρ of the annular ring does not exceed 90 percent of ρ maximum, thereby avoiding the discrepancies which may exist for large values of ρ , but which will not seriously effect the appearance of the spot diagrams. It is suggested that these annular curves are constructed for those pencils of interest with $\rho < 0.9\rho_{\text{max}}$. using points on the annular circle whose polar angles are 0, 45, 90, 135, 180°. This involves five ray traces, of which three are skew.

When the quality of agreement is being examined by using tangential traces, at least five rays for each pencil have to be traced in order to obtain the shape of the curve. Furthermore, if the angle criteria are being used, information at each surface is required for these rays. In view of the rather unsatisfactory nature of the angle criteria, and the fact that skew rays are not examined at all, the construction of the predicted and true annular curves along the lines suggested will give a better indication of the quality of agreement (for either secondary or tertiary predictions) than the tangential curves. The number of ray traces and predictions is the same in either case, and the ray trace programme will handle skew rays with equal facility as tangential rays so that no extra work is required for this method. It is, of course, realised that a single annular curve in any pencil does not in general indicate the shape or light distribution of the corresponding geometrical image patch. The method is intended only to check agreement, spot diagrams being the final arbiter in assessing the geometrical image-forming properties of the system.

PART IV - Application of Aberration Coefficients

15. Use of Coefficients.

The use of the aberration coefficients in optical design has been described in a general way in the paper by Cruickshank and Hills¹⁰. In their paper it was assumed that the first three orders of coefficients alone sufficiently well represent the state of correction of the optical system. It is the intention of this section of the thesis to indicate the methods presented in their paper and in addition present the means of extending the range of usefulness of the coefficients to those cases where it is known that the first three orders do not sufficiently well represent the state of correction of the system.

a) Good agreement between $\underline{\varepsilon}'[3]$ and $\underline{\varepsilon}'_k$.

The geometrical representation of the various types of aberrations can be most easily realised when the series for $\underline{\varepsilon}'_k$ is written in the polar coordinates described in Section 4. Writing (4.18) in full and including the secondary and tertiary coefficients, we get

$$\begin{aligned} \varepsilon'_{yk} = & [\sigma_1 \cos \varphi^3 + \sigma_2 (\cos 2\vartheta + 2) \rho^2 \bar{H} + (3\sigma_3 + \sigma_4) \cos \varphi \rho \bar{H}^2 + \sigma_5 \bar{H}^3] \\ & + [\mu_1 \cos \varphi^5 + (\mu_2 + \mu_3 \cos 2\vartheta) \rho^4 \bar{H} + (\mu_4 + \mu_6 \cos^2 \vartheta) \cos \varphi \rho^3 \bar{H}^2 + \\ & + (\mu_7 + \mu_8 \cos 2\vartheta) \rho^2 \bar{H}^3 + \mu_9 \cos \varphi \bar{H}^4 + \mu_{12} \bar{H}^5] + [\tau_1 \cos \varphi^7 + \\ & + (\tau_2 + \tau_3 \cos 2\vartheta) \rho^6 \bar{H} + (\tau_4 + \tau_6 \cos^2 \vartheta) \cos \varphi \rho^5 \bar{H}^2 + (\tau_7 + \tau_8 \cos 2\vartheta + \tau_{10} \cos 4\vartheta) \rho^4 \bar{H}^3 + \\ & + (\tau_{11} + \tau_{12} \cos^2 \vartheta) \cos \varphi \rho^3 \bar{H}^4 + (\tau_{15} + \tau_{16} \cos 2\vartheta) \rho^2 \bar{H}^5 + \tau_{18} \cos \varphi \bar{H}^6 + \tau_{20} \bar{H}^7] + O(9) \end{aligned}$$

$$\begin{aligned}
e'_{2k} = & [\sigma_1 \sin \rho^3 + \sigma_2 \sin 2 \rho \bar{H} + (\sigma_3 + \sigma_4) \sin \rho \bar{H}^2] + [\mu_1 \sin \rho^5 + \\
& + \mu_3 \sin 2 \rho \bar{H} + (\mu_5 + \mu_6 \cos^2 \theta) \sin \rho \bar{H}^2 + \mu_9 \sin 2 \rho \bar{H}^3 + \\
& + \mu_{11} \sin \rho \bar{H}^4] + [\tau_1 \sin \rho^7 + \tau_3 \sin 2 \rho \bar{H} + (\tau_5 + \tau_6 \cos^2 \theta) \sin \rho \bar{H}^2 \\
& + (\tau_9 \sin 2 \theta + \tau_{10} \sin 4 \theta) \rho^4 \bar{H}^3 + (\tau_{13} + \tau_{14} \cos^2 \theta) \sin \rho \bar{H}^4 + \\
& + \tau_{17} \sin 2 \rho \bar{H}^5 + \tau_{19} \sin \rho \bar{H}^6] + O(9)
\end{aligned} \tag{15.1}$$

The relationships between the σ, μ, τ and the final aberration coefficients are as follows, omitting the subscripts and primes.

$\sigma_1 = A$	$\mu_1 = S_1$	$\mu_7 = \frac{1}{2} (2\bar{S}_3 + \bar{S}_4 + S_5)$
$\sigma_2 = \bar{A}$	$\mu_2 = \bar{S}_1 + \frac{1}{2}S_2$	$\mu_8 = \frac{1}{2} (\bar{S}_4 + S_5)$
$\sigma_3 = \frac{1}{2}\bar{B}$	$\mu_3 = \frac{1}{2}S_2$	$\mu_9 = \frac{1}{2}S_5$
$\sigma_4 = C - \frac{1}{2}\bar{B}$	$\mu_4 = \bar{S}_2 + S_3$	$\mu_{10} = \bar{S}_5 + S_6$
$\sigma_5 = \bar{C}$	$\mu_5 = S_3$	$\mu_{11} = S_6$
	$\mu_6 = S_4$	$\mu_{12} = \bar{S}_6$

$\tau_1 = T_1$	$\tau_{11} = \bar{T}_5 + T_6$
$\tau_2 = \bar{T}_1 + \frac{1}{2}T_2$	$\tau_{12} = \bar{T}_7 + T_8$
$\tau_3 = \frac{1}{2}T_2$	$\tau_{13} = T_6$
$\tau_4 = \bar{T}_2 + T_3$	$\tau_{14} = T_8$
$\tau_5 = T_3$	$\tau_{15} = \bar{T}_6 + \frac{1}{2}\bar{T}_8 + T_9$
$\tau_6 = T_4$	$\tau_{16} = \frac{1}{2} (\bar{T}_8 + T_9)$
$\tau_7 = \bar{T}_3 + \frac{1}{2}\bar{T}_4 + \frac{1}{2}T_5 + \frac{3}{8}T_7$	$\tau_{17} = \frac{1}{2}T_9$
$\tau_8 = \frac{1}{2} (\bar{T}_4 + T_5 + T_7)$	$\tau_{18} = \bar{T}_9 + T_{10}$
$\tau_9 = \frac{1}{2}T_5 + \frac{1}{4}T_7$	$\tau_{19} = T_{10}$
$\tau_{10} = \frac{1}{8}T_7$	$\tau_{20} = \bar{T}_{10}$

Note that the factor $1/N'_k v'_{sk}$ has been included in the coefficients A, S, T.

The geometrical representation of the various types of aberrations are given in fig. 18 (Fig. 3 in the Cruickshank and Hills paper).

The presentation of ε_k^i in the form (15.1) is of considerable assistance in deciding how to balance the aberrations of the various orders in order to achieve a satisfactory final image. A good illustration of the procedure adopted is given in Section 7 of ref. 10, where the various stages of the correction of the telephoto system used in this thesis are shown. The power of the aberration coefficients method lies not so much in the representation of ε_k^i as a power series in the initial coordinates, but in the fact that each final coefficient is the linear sum of the contributions by each surface of the system, together with the fact that the coefficients characterise systems of rays. As can be seen in Section 5, the coefficients are independent of the values of Y_1 , V_1 , they depend only on the constitution of the system. Resulting from these properties of the coefficients, the designer has at his disposal a powerful tool for use in correcting optical systems.

From the contributions to the final aberration coefficients by the various surfaces of an optical system, one can gain insight into the "raison d'être" of the system. As an example, table VII gives the contributions to the coefficients for the Vega. The final coefficients are under the column marked Σ . Consider the spherical aberration of this system. The third, fifth and seventh order coefficients controlling this aberration are σ_1 , μ_1 , τ_1 . It can be seen that the final coefficients in all three orders are negative. This is the result of the large negative contributions by the fourth and fifth surfaces. These contributions are to some extent compensated by the first and eighth surfaces. In comparison, the remaining surfaces contribute little to this aberration. On the other hand, the contributions to primary coma, σ_2 , are fairly well distributed amongst the surfaces. μ_2 and μ_3 control secondary circular coma, and the surface contributing most to this aberration is the first. By scrutinising the surface contributions to the coefficients in this manner, one gains an understanding of why the system behaves as it does.

So far, the aberration coefficients have been considered only as a means of describing the behaviour of a system. Now, when the coefficients are to be used as an aid in the correction of a system, it is of considerable advantage to have at least the first derivatives of the final coefficients with respect to the parameters

Table VII

	1	2	3	4	5	6	7	8	Σ
σ_1	30.59	0.2563	5.814	-34.24	-42.52	13.53	-1.837	27.65	-0.7633
σ_2	1.173	-0.1094	-0.4827	0.7130	-0.8675	1.037	-0.5494	-1.102	-0.1881
σ_3	0.0450	0.0467	0.0401	-0.0148	-0.0177	0.0795	-0.1644	0.0439	0.0583
σ_4	1.211	-0.8229	1.689	-2.089	-2.089	1.689	-0.8229	1.211	-0.0240
σ_5	0.0481	0.3313	-0.1436	0.0438	-0.0430	0.1356	-0.2953	-0.0500	0.0270
μ_1	714.1	-9.565	48.57	-983.1	-1155	94.08	42.54	587.5	-661.2
μ_2	119.2	-20.26	53.77	-31.07	-28.20	-27.88	32.73	-66.00	32.24
μ_3	91.79	-22.99	55.38	-42.19	11.15	-41.21	26.10	-56.47	21.55
μ_4	30.82	26.33	20.27	-60.67	-57.16	23.45	13.28	21.67	17.99
μ_5	23.78	5.368	27.45	-57.29	-50.78	24.19	1.913	18.90	-6.464
μ_6	-20.71	-3.694	-3.744	42.07	52.74	-11.86	-4.815	-24.73	25.26
μ_7	1.774	2.225	2.078	-3.259	3.897	-0.5308	-3.612	-3.048	-0.4778
μ_8	0.8624	0.4950	3.972	-4.946	4.742	-3.176	-0.7778	-1.373	-0.2010
μ_9	1.259	-0.2142	4.018	-4.623	4.330	-3.077	0.1228	-1.891	-0.0753
μ_{10}	-0.1053	-0.0733	0.1965	-0.4637	-0.3195	0.2572	-0.1681	0.1722	-0.5040
μ_{11}	-0.2018	0.2559	0.0880	-0.2719	-0.2294	0.1237	0.1599	-0.0736	-0.1493
μ_{12}	-0.0077	0.0666	-0.1073	0.0713	0.0435	0.0083	-0.0217	-0.0675	-0.0145
τ_1	19660	-901.5	1939	-31830	-35940	-836.9	1420	16020	-30470
τ_2	4455	5.147	1992	-268.1	539.8	-3266	1841	-4260	1039
τ_3	3702	-385.5	2090	-653.5	1848	-3310	1544	-4008	826.1
τ_4	943.3	806.2	390.4	-2182	-1844	-42.73	1321	746.9	138.8
τ_5	659.5	203.5	661.5	-2019	-1700	365.5	402.3	541.7	-884.8
τ_6	-460.2	-159.4	715.6	1553	2382	300.4	-250.4	-654.5	3426
τ_7	-3.508	54.92	87.37	-162.3	102.1	45.99	-119.4	-9.344	-4.244
τ_8	-46.79	40.03	111.2	-209.1	137.0	12.13	-112.6	45.12	-22.93
τ_9	-1.966	3.576	140.3	-191.9	129.5	-36.10	-38.94	-9.613	-5.201
τ_{10}	-18.00	4.464	-10.21	1.959	-13.48	23.35	-16.45	23.37	-5.007
τ_{11}	-0.047	25.76	-7.835	-22.04	-25.89	17.98	-17.99	11.50	-18.55
τ_{12}	4.311	-28.26	22.28	-28.16	-29.77	33.00	-54.32	13.48	-67.44
τ_{13}	-2.656	5.485	4.405	-14.86	-17.41	9.354	1.909	4.700	-9.070
τ_{14}	9.832	-8.906	17.11	-34.96	-35.24	22.87	-20.18	20.11	-29.35
τ_{15}	-0.5829	2.534	-3.217	2.571	-1.727	2.202	-2.854	0.2224	-0.8515
τ_{16}	-0.4811	0.2023	-1.147	1.229	-1.469	0.1088	0.4344	0.4647	-0.6579
τ_{17}	-0.6695	0.5606	-2.177	2.389	-0.9879	0.1832	0.3124	0.2310	-0.1578
τ_{18}	0.0340	-0.0442	0.1304	-0.0614	-0.9613	0.5934	0.5816	0.3864	0.6588
τ_{19}	0.0853	-0.0201	-0.1224	0.1524	-0.3337	0.2506	-0.0068	0.2044	0.2097
τ_{20}	0.0033	0.0166	-0.0048	-0.0070	-0.0157	0.0287	-0.0283	0.0232	0.0159

SPHERICAL ABERRATION S TYPE $p_{3.5,7}$		ASYMMETRIC ABERRATIONS C TYPES $R_{1,3,5}$			ASTIGMATIC ABERRATIONS S TYPES $p_{1,3,5}$		DISTORTION C TYPE $H_{3,5,7}$	
THIRD ORDER		LINEAR COMA $p_{2,4,6} R^1$	CUBIC COMA $p_{2,4} R^3$	QUINTIC COMA $p^2 R^5$	LINEAR ASTIGMATISM $p^1 R_{2,4,6}$	CUBIC ASTIGMATISM $p^3 R_{2,4}$	QUINTIC ASTIGMATISM $p^5 R_2$	THIRD ORDER
FIFTH ORDER								FIFTH ORDER
SEVENTH ORDER								SEVENTH ORDER

Fig. 18 Geometrical representation of aberrations

of the system. With these derivatives it is then possible to estimate the effect of moderate changes of construction, not only of single changes, but simultaneous changes of a number of parameters, which leads to considerable economy of design. Care must be taken when assessing the effect of changes since the focal length is usually altered as well. Before a direct comparison can be made the focal lengths of the original and modified systems must be identical. If the original system has unit focal length, then the multiplication of the new ε'_k or $\underline{\varepsilon}$ [3] by the new v'_{ak} will make these displacements also correspond to a system of unit focal length.

It has been found, using only the first derivatives of the coefficients, that changes of parameters not exceeding about 5 percent (moderate changes) can be made with reasonable accuracy. If changes greater than this are required then it would be desirable to recompute the coefficients and derivatives after, say, each 10 percent change. The use of second derivatives would increase the accuracy of the estimate of the new coefficients, but, in view of the labour required to obtain these derivatives, their use is not considered worth while. In any case, the derivatives are used only as a guide to the required changes, the performance of the system being more accurately assessed from the actual coefficients computed after a change is made. These techniques are discussed in some detail in the paper by Cruickshank and Hills. Methods of obtaining the first derivatives are discussed later.

If a parameter of a system which precedes the diaphragm is altered, in general the value of p also is changed. Since $\bar{O}T$ paracanon-ical coordinates depend on p , strictly speaking, each time the parameters are so altered, a new set of $\bar{O}T$ coordinates should be computed corresponding to the new value of p . However, if the changes are moderate, the effect of the change in p is usually negligible compared with the change in the aberration coefficients so that p can be considered constant and the coordinates need not be altered after such modifications of the system. When the same value of p is used throughout, changes of parameters before the diaphragm necessarily imply a change in the position of the diaphragm, which will not be serious unless large changes are made. Note also that a change in scale of the system (to readjust the focal length perhaps) changes p and l_0 , if the latter is not infinite. If a large change is made with p constant, it may be found that the diaphragm plane has been moved into the glass of a component.

By way of example, changes of curvature of three surfaces were made in each of the systems, the Vega and the Sonnar. The changes made were, for the Vega $c_2 + 4.6\%$, $c_4 - 0.1\%$, $c_6 + 2.0\%$ and for the Sonnar $c_1 - 2.0\%$, $c_3 + 1.1\%$, $c_5 - 5.2\%$. The predicted and computed coefficients for the modified systems, as well as those for the original systems are shown in Table VIII. Note here that the ordinary algebraic coefficients have been tabulated since these are the ones used by the programmes. The geometrical interpretations of the various types of aberrations are not important in this example.

On the whole, the predicted coefficients are satisfactory. Those cases where the agreement is not so good, mainly in the tertiary coefficients, generally correspond to large changes in the coefficients, e.g. \bar{T}_1 and T_7 for the Vega. The T_7 predicted coefficient indicates a 93 percent change, whereas the actual change is about 80 percent. However, since the tertiary coefficients multiply seventh powers of the coordinates of a ray, an occasional error of this magnitude will not have a serious effect when the coefficient is comparatively small.

b) Bad agreement between $\varepsilon'_1[3]$ and ε'_k

In those cases where $\varepsilon'_1[3]$ does not closely approximate ε'_k , it is still possible to use the coefficients as an aid in the understanding and correction of a system. Now, although the first three orders of coefficients and their contributions alone do not exactly represent the aberrations of the system, they are still capable of indicating the surfaces whose contributions dominate the final amount of any selected aberration. For instance, in the notes on the Sonnar in Section 13., it was mentioned that the true tangential curve could be closely approximated by calculating the effects of higher order spherical aberration at those surfaces whose primary, secondary and tertiary contributions were predominant. Admittedly, an approximation to only higher order spherical aberration can be computed so far, the point is that the surfaces at which the higher order contributions are likely to be predominant can be assessed with reasonable confidence on the basis of the first three orders. In the Vega, Table VII, the contributions to circular coma $\sigma_2; \mu_2, \mu_3$ and τ_2, τ_3 are comparatively large and are all positive at the first surface. It is thus likely that this state of affairs will continue in the higher orders. However, the eighth surface contributes negatively to this aberration, the contributions here being of sufficient magnitude to just about cancel the effect of the first surface. There is no reason to suppose that this cancellation will not continue in the higher order contributions to circular coma.

Table VIII

Vega				Sonhar		
	Original	Predicted	Computed	Original	Predicted	Computed
A	-0.7633	2.292	2.519	0.1778	0.1716	0.1742
\bar{A}	-0.1881	-0.4272	-0.4148	0.0242	0.0054	0.0065
\bar{B}	0.1165	0.0360	0.0389	-0.0727	-0.1155	-0.1148
C	0.0342	0.0012	0.0027	0.1205	0.0989	0.0993
\bar{C}	0.0270	0.0026	0.0034	0.2838	0.2619	0.2619
S_1	-661.2	-560.8	-546.9	-1.389	-1.482	-1.468
\bar{S}_1	10.69	3.185	3.905	-0.4599	-0.5539	-0.5523
S_2	43.11	12.50	16.04	-1.756	-2.123	-2.115
\bar{S}_2	24.46	20.51	20.40	-1.062	-1.273	-1.275
S_3	-6.464	-6.538	-6.513	-1.302	-1.434	-1.434
\bar{S}_3	-0.2767	-0.5226	-0.5115	0.0728	-0.0032	-0.0034
S_4	25.26	18.06	17.87	-1.328	-1.532	-1.537
\bar{S}_4	-0.2515	0.1489	0.1107	-1.063	-1.157	-1.159
S_5	-0.1506	-0.1672	-0.1875	0.0898	-0.0149	-0.0184
\bar{S}_5	-0.3547	-0.2669	-0.2646	-0.0809	-0.1483	-0.1490
S_6	-0.1493	-0.1150	-0.1160	-0.3708	-0.3801	-0.3804
\bar{S}_6	-0.0145	-0.0110	-0.0102	-0.2756	-0.2937	-0.2936
T_1	-30475	-27999	-27410	-8.233	-9.530	-8.935
\bar{T}_1	213.1	15.97	46.04	-1.953	-2.352	-2.361
T_2	1652	-86.98	164.7	-12.63	-15.16	-15.19
\bar{T}_2	1024	902.5	898.2	-2.384	-3.227	-3.271
T_3	-884.8	-854.5	-848.1	-7.220	-7.980	-7.997
\bar{T}_3	13.68	1.984	2.685	0.6395	0.3704	0.3605
T_4	3426	2909	2884	-0.6876	-2.195	-2.319
\bar{T}_4	-15.43	5.113	1.588	1.434	1.220	1.176
T_5	9.625	-1.574	-4.143	4.263	3.686	3.627
\bar{T}_5	-9.484	-8.708	-8.622	1.896	1.805	1.787
T_6	-9.071	-7.775	-7.907	-0.8988	-0.8738	-0.8828
\bar{T}_6	-0.1937	-0.1032	-0.0926	0.4851	0.4863	0.4848
T_7	-40.06	-2.908	-8.253	2.547	2.692	2.654
\bar{T}_7	-38.09	-32.24	-32.18	2.481	2.750	2.744
T_8	-29.35	-19.97	-19.93	5.994	6.332	6.315
\bar{T}_8	-1.000	-0.7817	-0.7805	0.3611	0.6641	0.6649
T_9	-0.3155	-0.0581	-0.0413	0.6720	0.8926	0.8924
\bar{T}_9	0.4491	0.4014	0.3959	1.403	1.559	1.562
T_{10}	0.2098	0.1828	0.1851	-0.6713	-0.6168	-0.6162
\bar{T}_{10}	0.0159	0.0134	0.0132	-0.3938	-0.3580	-0.3569

It must be remembered that any inaccuracy in agreement between predicted and true displacements is not a result of inaccurate coefficients, but simply that the series for ε'_k has been terminated too soon. The first three orders of coefficients work quite well for moderate values of \bar{H} and ρ .

In a reasonably well corrected system, that is, after the initial design, the primary aberrations tend to be well distributed amongst the various surfaces, resulting in small final aberrations. The distribution of the secondary aberrations is generally such that the contributions do not balance one another quite so well. In the tertiary aberrations, it is usual to find that the final coefficients are of the same order of magnitude as the contributions as a result of a marked lack of balance amongst the contributions. This can be seen in Table VII. Under these conditions, a small change in a parameter will have a marked effect on the primary aberrations, a less effect on the secondary aberrations and little effect on the tertiary aberrations. This can be seen in Table IX which gives the first curvature derivatives for a few surfaces of the Vega. Note here that the coefficients are A , S , T and not σ , μ , τ . In this table it will be noticed that the ratio of derivative/coefficient generally decreases as the order of the coefficient increase. For example, the coefficients of distortion \bar{C} , \bar{S}_4 , \bar{T}_{10} and their derivatives indicate this quite well. In the case of spherical aberration A , S_4 , T , the effect is most noticeable between the primary and secondary coefficients. The tertiary ratio at the surfaces shown is about the same as the secondary ratio, the ratio of $\frac{dT}{dC}/T$, having a maximum of approximately $|3|$. Thus, at the first surface, a change in curvature of 1 percent would cause a 36 percent change in A , a 2 percent change in S , and a 3 percent change in T . It is for this reason then that the higher order aberrations can be considered comparatively stable under moderate changes of the parameters of a system.

This has the undesirable effect of making it practically impossible to vary the higher order aberrations without seriously upsetting the balance of the lower order aberrations once the initial design has been made. On the other hand, this property of the coefficients enables the designer to satisfactorily determine the effects of moderate changes in those regions where it is known that the first three orders of coefficients do not accurately represent the aberrations of the system.

Denote the difference $\varepsilon'_k - \varepsilon'[3]$ by Δ . There is, of course, a particular pair of values of Δ corresponding to each value of

ρ , e and \bar{H} used in (15.1). Δ represents the effect of the 9th and higher order aberrations, thus for moderate changes of the system Δ can be considered to be constant.

Thus $\delta(\epsilon_k') \approx \delta(\epsilon'[3])$. Now $\delta(\epsilon'[3])$ can be satisfactorily determined with the use of the first derivatives (Section 15(a)), hence allowing for the appropriate values of Δ , ϵ_k' for the modified system can be reliably predicted. Initially, for those pencils for which the tertiary agreement is poor, one requires sets of ray traces, which give the values of ϵ_k' , and corresponding sets of $\epsilon'[3]$. The difference between the corresponding pairs then of course gives Δ_y, Δ_z for the particular rays chosen. Having once obtained the Δ in this manner, recourse to further ray traces should only be necessary to check the final performance unless the changes made in the system are large.

To illustrate this method, the effect of the changes to the Vega and Sonnar listed in Section 15(a) are shown in Figs. 19, 20. In these diagrams, the tangential curves for the original systems are given in the columns so marked, the true curves being in solid line, the broken line indicating $\epsilon_y'[3]$. The required values of Δ (in the case of tangential pencils, Δ_y only) are the separations of the curves. The changes made in these systems were chosen mainly to cause significant changes in the tangential curves, although some attempt was made to reduce distortion (which unfortunately introduced more coma, particularly in the case of the Vega). In the diagrams for the modified (new) systems, the broken line indicates $\epsilon_y'[3] + \Delta_y$, where $\epsilon_y'[3]$ was obtained from the actual coefficients computed for the modified systems. As mentioned in Section 15 (a), the coefficients predicted with derivatives are used only as a guide to the required changes.

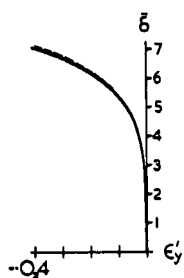
Considering the large magnitudes of Δ_y in some cases, it can be seen that the true curves have been predicted remarkably well for all pencils in both systems.

Table IX

	Coefficient	d/dc_1	d/dc_2	d/dc_3	d/dc_4
A	-0.7633	27.50	-7.132	-24.99	21.34
\bar{A}	-0.1881	2.131	-0.2046	-0.6558	-1.058
\bar{B}	0.1165	0.3521	-0.2573	0.1562	-0.1346
C	0.0342	0.03658	-0.3184	-0.1359	0.1466
\bar{C}	0.0270	0.0981	-0.1058	-0.0425	-0.0179
S_1	-661.2	1591	-497.7	-1252	892.3
\bar{S}_1	10.69	93.24	-32.92	-21.27	-5.406
S_2	43.11	362.4	-129.5	-76.32	-26.37
\bar{S}_2	24.46	6.935	-4.702	15.27	-12.62
S_3	-6.464	15.53	-9.610	-9.560	8.240
\bar{S}_3	-0.2767	0.6326	-0.023	0.1706	-1.240
S_4	25.26	-22.10	2.828	41.65	-35.29
\bar{S}_4	-0.2515	-3.413	0.6128	1.869	1.495
S_5	-0.1506	-3.053	0.2181	1.662	-0.2312
\bar{S}_5	-0.3547	-0.1279	0.1114	-0.1377	0.3000
S_6	-0.1493	-0.1343	0.0957	0.0111	0.0672
\bar{S}_6	-0.0145	-0.0208	0.0151	0.0087	0.0017
T_1	-30475	80483	-27366	-54902	34759
\bar{T}_1	213.1	3980	-1531.4	-970.2	498.5
T_2	1652	21466	-8681	-2224	-133.5
\bar{T}_2	1024	101.0	-30.22	751.8	-465.5
T_3	-884.9	836.8	-526.1	694.8	574.9
\bar{T}_3	13.68	24.61	-19.22	13.64	-38.40
T_4	3426	-2803	701.8	4032	-2854
\bar{T}_4	-15.43	-296.0	116.5	106.6	-8.041
T_5	9.625	-275.1	63.54	125.9	-113.4
\bar{T}_5	-9.484	-0.7307	-2.772	-7.582	5.577
T_6	-9.071	-8.992	3.450	1.701	2.491
\bar{T}_6	-0.1937	0.3591	-0.2500	-0.4431	0.6720
T_7	-40.06	-432.8	158.1	109.3	32.88
\bar{T}_7	-38.09	-5.963	9.755	-22.05	15.85
T_8	-29.35	1.685	8.816	-39.54	34.52
\bar{T}_8	-1.000	4.135	-0.6390	-3.811	1.479
T_9	-0.3155	3.490	-0.3780	-2.611	1.530
\bar{T}_9	0.4491	0.1317	-0.0089	0.0076	-0.1973
T_{10}	0.2098	0.1087	-0.0642	-0.0325	-0.0577
\bar{T}_{10}	0.0159	-0.0020	0.0034	0.0002	-0.0150

ORIGINAL

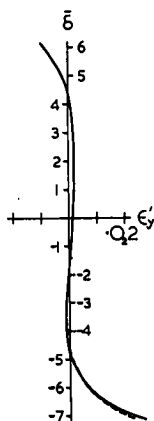
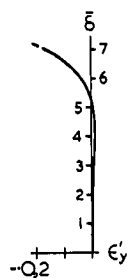
--- $\epsilon'[3]$



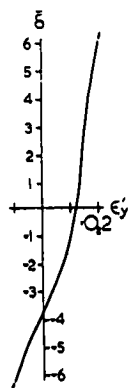
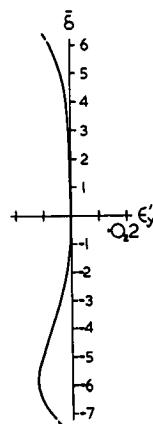
0°

NEW

--- $\epsilon'[3] + \Delta$



10°



20°

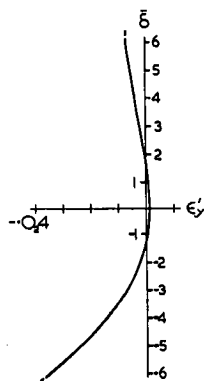
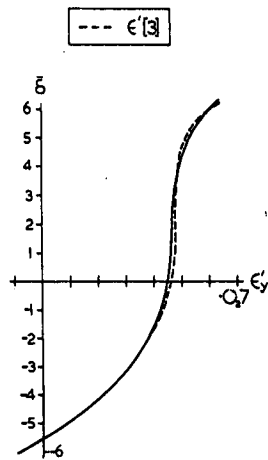


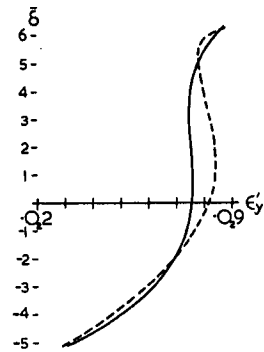
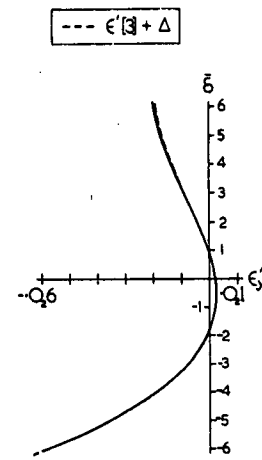
Fig. 19 Vega

ORIGINAL

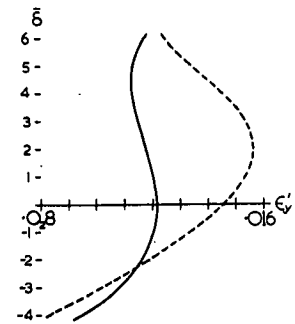
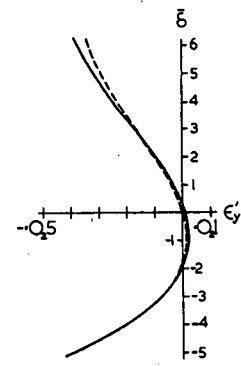
NEW



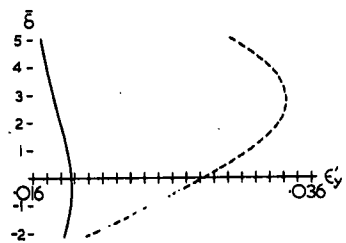
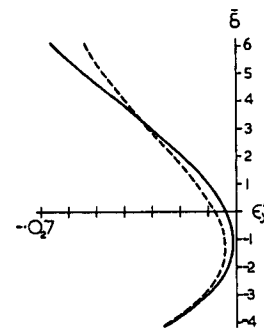
30°



35°



40°



45°

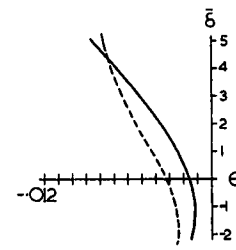


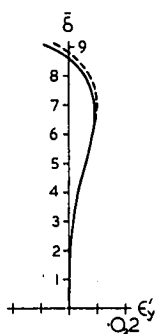
Fig. 19 Vega

ORIGINAL

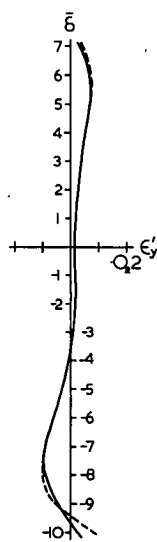
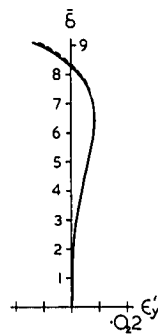
NEW

--- $\epsilon'[3]$

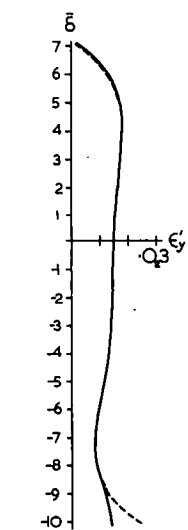
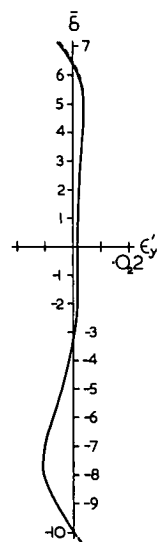
--- $\epsilon'[3] + \Delta$



0°



5°



10°

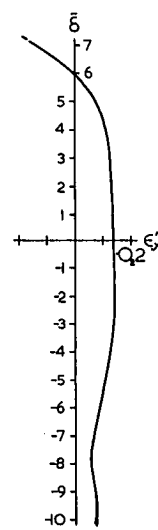
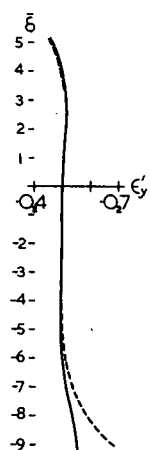


Fig. 20 Sonnar

ORIGINAL

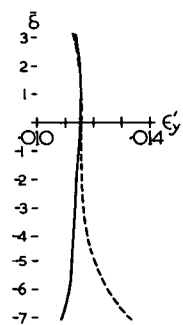
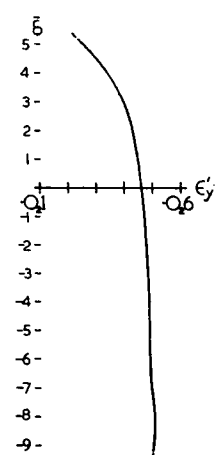
--- $\epsilon'_{[3]}$



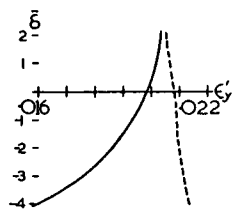
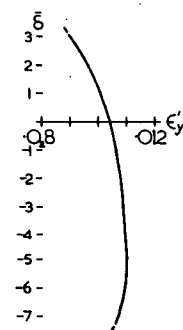
15°

NEW

--- $\epsilon'_{[3]} + \Delta$



20°



25°

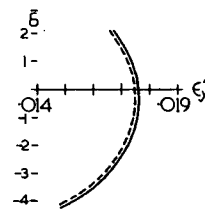


Fig. 20 Sonnar

16. Derivatives of the Coefficients.

The theoretical methods for obtaining the derivatives of the coefficients with respect to the various parameters of a system are given in M Section 120 - 162. In particular, the first separation derivatives of the primary aberration coefficients can be computed from M 128.51-54, the first curvature derivatives for the same coefficients from M 142.4 - 7, and the first chromatic derivatives for the primary coefficients from M 152.5 - 8. However, to obtain explicit expressions for the first derivatives of all the secondary and tertiary coefficients would involve considerable work. As a result, only the first derivatives of the primary coefficients have been available so far for use in the routine correction of optical systems.

Now that a comparatively fast programme for the computation of the first three orders of coefficients has been developed, the first derivatives of these coefficients can also be obtained by numerical means. Given the specification and coefficients of any system, the derivatives of the coefficients with respect to any particular parameter can be found simply by making a small change in the desired parameter and recomputing the coefficients. Division of the difference between the two set of coefficients by the increment in the parameter then gives the desired first derivatives. The process is of course repeated for each parameter since only one of these must be different from normal during any one computation.

The magnitude of the change to be made in any one parameter is governed by two conditions. Firstly, it must not be too small or there will be no significant figures in the difference between the coefficients, and secondly, it must not be too large or the results will be influenced by the second derivatives. In order to determine the optimum increment, five sets of separation and curvature derivatives were computed for a cemented doublet, one for each of five different incremental changes of the parameters. The percentage change in the parameters were 0.0001, 0.001, 0.01, 0.1 and 1.0. Owing to the fact that at present only the derivatives of the primary coefficients can be computed from theoretical expressions, the derivatives of the secondary and tertiary coefficients were checked by means of the "identities of homogeneity" M Section 123. The identities are formed by considering the effect of a change of scale on an optical system.

In M Section 123 it is shown that, for any aberration coefficient G ,

$$\theta G - l_w \frac{\partial G}{\partial l_{01}} - p \frac{\partial G}{\partial p} = \Phi G \quad (16.1)$$

where θ is the differential operator

$$\theta = \sum_{i=1}^k c_j \frac{\partial}{\partial c_j} - \sum_{j=2}^k d_j \frac{\partial}{\partial d_j} \quad (16.2)$$

and Φ is factor which depends on the order n and the μ, ν of the particular coefficient.

If the object is at infinity, then the term containing l_{01} is absent from (16.1). Accordingly, the object distance for the doublet was made infinite to simplify the work.

The derivatives of the coefficients with respect to p which occur in (16.1) were obtained from M Section 34. In this section, the changes in the primary and secondary "a" coefficients are given for a finite change χ in p , (M34.41, 43) using $\bar{O}T$ coordinates. By letting χ tend to zero, the derivatives of the coefficients with respect to p can be found. An extension of the work of M Section 34 yields the required derivatives of the tertiary coefficients.

Table X lists the identities for the first three orders of augmented aberration coefficients (see 3.18). In this table the identity for the most important paraxial coefficient v'_{ak} is given. The primes and the subscripts ak have been omitted from the aberration coefficients.

Applying these identities to the results from the cemented doublet, it was found that the 0.01 percent change in parameters gave the most accurate values for the derivatives. That this percentage change is the best was corroborated by the work done on the Vega and Sonnar. The curvature derivatives of the primary coefficients used in the example to predict the effect of changes closely agreed with those computed from the theoretical expressions of M 142.4 - 7. Table XI shows the magnitude of the identities of Table X for the doublet, the "c" and "d" derivatives in θ (16.2) having been found by the numerical means mentioned earlier.

Table X

OT Coordinates, object at infinity.

$\theta v'_{ak}$	$= v'_{ak}$		
θA	$= 3A$	θT_1	$= 7T_1$
$\theta \bar{A} - pA$	$= 2\bar{A}$	$\theta \bar{T}_1 - pT_1$	$= 6\bar{T}_1$
$\theta \bar{B} - 4p\bar{A}$	$= \bar{B}$	$\theta T_2 - 6pT_1$	$= 6T_2$
$\theta C - 2p\bar{A}$	$= C$	$\theta \bar{T}_2 - p(6\bar{T}_1 + T_2)$	$= 5\bar{T}_2$
$\theta \bar{C} - p(\bar{B} + C)$	$= 0$	$\theta T_3 - pT_2$	$= 5T_3$
		$\theta \bar{T}_3 - p(\bar{T}_2 + T_3)$	$= 4\bar{T}_3$
θS_1	$= 5S_1$	$\theta T_4 - 4pT_2$	$= 5T_4$
$\theta \bar{S}_1 - pS_1$	$= 4\bar{S}_1$	$\theta \bar{T}_4 - p(4\bar{T}_2 + T_4)$	$= 4\bar{T}_4$
$\theta S_2 - 4pS_1$	$= 4S_2$	$\theta T_5 - 2p(2T_3 + T_4)$	$= 4T_5$
$\theta \bar{S}_2 - p(4\bar{S}_1 + S_2)$	$= 3\bar{S}_2$	$\theta \bar{T}_5 - p(4\bar{T}_3 + 2\bar{T}_4 + T_5)$	$= 3\bar{T}_5$
$\theta S_3 - pS_2$	$= 3S_3$	$\theta T_6 - pT_5$	$= 3T_6$
$\theta \bar{S}_3 - p(\bar{S}_2 + S_3)$	$= 2\bar{S}_3$	$\theta \bar{T}_6 - p(\bar{T}_5 + T_6)$	$= 2\bar{T}_6$
$\theta S_4 - 2pS_2$	$= 3S_4$	$\theta T_7 - 2pT_4$	$= 4T_7$
$\theta \bar{S}_4 - p(2\bar{S}_2 + S_4)$	$= 2\bar{S}_4$	$\theta \bar{T}_7 - p(2\bar{T}_4 + T_7)$	$= 3\bar{T}_7$
$\theta S_5 - 2p(S_3 + S_4)$	$= 2S_5$	$\theta T_8 - p(2T_5 + 3T_7)$	$= 3T_8$
$\theta \bar{S}_5 - p(2\bar{S}_3 + 2\bar{S}_4 + S_5)$	$= \bar{S}_5$	$\theta \bar{T}_8 - p(2\bar{T}_5 + 3\bar{T}_7 + T_8)$	$= 2\bar{T}_8$
$\theta S_6 - pS_5$	$= S_6$	$\theta T_9 - 2p(T_6 + T_8)$	$= 2T_9$
$\theta \bar{S}_6 - p(\bar{S}_5 + S_6)$	$= 0$	$\theta \bar{T}_9 - p(2\bar{T}_6 + 2\bar{T}_8 + T_9)$	$= \bar{T}_9$
		$\theta T_{10} - pT_9$	$= T_{10}$
		$\theta \bar{T}_{10} - p(\bar{T}_9 + T_{10})$	$= 0$

Table XI

0.01% change in parameters c, d.

G	$\Theta G - p \frac{\partial G}{\partial p}$	ΦG	G	$\Theta G - p \frac{\partial G}{\partial p}$	ΦG
A	0.0768	0.0768	T ₁	-756.4	-756.3
\bar{A}	0.00706	0.00707	\bar{T}_1	-120.9	-120.9
\bar{B}	0.9897	0.9896	T ₂	-549.1	-549.0
C	0.8350	0.8350	\bar{T}_2	-105.6	-105.6
\bar{C}	-0.0 ₆ 4	0.0	T ₃	-56.89	-56.89
			\bar{T}_3	-12.69	-12.69
S ₁	-44.87	-44.87	T ₄	-128.5	-128.5
\bar{S}_1	-6.687	-6.692	\bar{T}_4	-36.37	-36.37
S ₂	-26.90	-26.90	T ₅	-18.68	-18.68
\bar{S}_2	-0.5124	-0.5126	\bar{T}_5	2.191	2.190
S ₃	1.978	1.978	T ₆	0.00559	0.00560
\bar{S}_3	-0.9073	-0.9073	\bar{T}_6	-0.7721	-0.7721
S ₄	-0.5634	-0.5635	T ₇	-7.619	-7.619
\bar{S}_4	1.0204	1.0203	\bar{T}_7	-0.4047	-0.4050
S ₅	2.6131	2.6128	T ₈	4.995	4.994
\bar{S}_5	0.8184	0.8184	\bar{T}_8	4.241	4.240
S ₆	-0.1182	-0.1182	T ₉	1.474	1.475
\bar{S}_6	0.0 ₆ 2	0.0	\bar{T}_9	-0.1242	-0.1243
			T ₁₀	0.03109	0.03109
			\bar{T}_{10}	-0.0 ₄ 2	0.0

It is rather wasteful of machine time to compute a complete set of curvature and separation derivatives using the aberration coefficients programme. On the DEUCE, the derivatives for a ten surface system would take $2\frac{1}{2}$ hours. However, as mentioned earlier, in any change of parameters, the derivatives of the primary coefficients have the greatest effect, and the hand computation of these can be easily done from the expressions in M Section 120 - 162. An examination of the derivatives so obtained will then indicate the surfaces and parameters that will be most effective when changed. The machine can then be used to obtain the derivatives of the higher order coefficients for the selected surfaces and parameters. In this way machine time can be kept to a minimum without sacrificing any of the power available to the designer that ensues from the use of coefficients and their derivatives.

APPENDIX 1

Final Aberration Coefficients of Tessar

+ 8.9 1 8 9 4 1.0 9	- 0 1
+ 1.4 5 1 3 6 7.1 9	- 0 1
- 4.3 0 1 0 8 2.0 2	- 0 2
+ 1.0 1 6 6 7 8.0 1	- 0 1
+ 4.8 2 8 1 6 9.2 6	- 0 3

- 7.7 2 4 2 9 7.8 6	+ 0 1
- 4.9 2 3 2 6 3.2 8	+ 0 0
- 1.9 3 1 2 3 6.1 1	+ 0 1
+ 1.0 6 3 1 3 1.0 3	+ 0 0
- 3.6 0 2 3 6 3.1 8	+ 0 0
- 6.0 1 5 8 7 2.9 7	- 0 1
+ 2.2 7 8 8 9 3.8 1	- 0 1
- 1.6 4 5 5 2 9.4 0	+ 0 0
- 1.4 7 9 9 5 1.2 1	+ 0 0
- 2.8 3 4 5 4 5.4 7	- 0 1
- 5.7 0 7 7 7 3.4 8	- 0 1
- 1.9 3 6 9 1 7.0 1	- 0 2

- 2.0 4 4 3 5 6.9 0	+ 0 3
- 1.3 3 7 3 4 0.5 8	+ 0 3
- 1.4 6 9 1 8 4.6 1	+ 0 2
+ 2.3 8 1 1 5 4.3 6	+ 0 2
- 6.3 4 4 1 6 0.7 9	+ 0 1
- 4.0 6 6 7 5 1.7 0	+ 0 0
- 1.0 7 1 9 7 8.0 8	+ 0 1
- 1.1 2 3 1 9 9.5 7	+ 0 1
- 1.9 9 2 0 4 0.0 8	+ 0 0
+ 2.7 0 8 5 9 4.1 6	- 0 2

- 2.1 4 0 6 2 5.7 4	+ 0 2
+ 4.2 4 6 7 3 3.9 4	+ 0 1
- 1.9 9 7 4 6 8.9 3	+ 0 1
- 3.6 0 7 1 9 0.7 8	+ 0 1
- 1.3 6 7 5 3 1.7 8	+ 0 1
- 1.6 0 7 2 4 2.6 7	+ 0 0
+ 3.0 5 9 7 5 3.6 3	+ 0 0
- 4.5 6 5 4 3 0.2 8	+ 0 0
+ 1.1 0 7 3 9 9.9 0	+ 0 0
+ 7.6 0 4 0 0 4.7 6	- 0 2

APPENDIX 2

Portion of Biotar 12° Ray Trace Results

ϵ'_{yk}	ϵ'_{zk}	V'_k	W'_k	y	z
- 002680	+ 000000	+ 00.140423	+ 00.000000	+ 00	+ 00
- 002690	+ 000110	+ 00.140509	+ 00.039185	+ 00	+ 01
- 002721	+ 000223	+ 00.140772	+ 00.078565	+ 00	+ 02
- 002771	+ 000339	+ 00.141234	+ 00.118334	+ 00	+ 03
- 002837	+ 000437	+ 00.141934	+ 00.158684	+ 00	+ 04
- 002914	+ 000469	+ 00.142932	+ 00.199795	+ 00	+ 05
- 002992	+ 000331	+ 00.144321	+ 00.241818	+ 00	+ 06
- 003054	- 000185	+ 00.146236	+ 00.284825	+ 00	+ 07
- 003066	- 001504	+ 00.148884	+ 00.328675	+ 00	+ 08
- 002725	+ 000000	+ 00.179930	+ 00.000000	+ 01	+ 00
- 002738	+ 000085	+ 00.180050	+ 00.039436	+ 01	+ 01
- 002781	+ 000176	+ 00.180415	+ 00.079076	+ 01	+ 02
- 002854	+ 000266	+ 00.181047	+ 00.119124	+ 01	+ 03
- 002963	+ 000339	+ 00.181985	+ 00.159784	+ 01	+ 04
- 003112	+ 000346	+ 00.183288	+ 00.201257	+ 01	+ 05
- 003313	+ 000181	+ 00.185043	+ 00.243723	+ 01	+ 06
- 003582	- 000363	+ 00.187374	+ 00.287303	+ 01	+ 07
- 003951	- 001726	+ 00.190459	+ 00.331940	+ 01	+ 08
- 002859	+ 000000	+ 00.220213	+ 00.000000	+ 02	+ 00
- 002878	+ 000054	+ 00.220372	+ 00.039752	+ 02	+ 01
- 002935	+ 000109	+ 00.220853	+ 00.079717	+ 02	+ 02
- 003038	+ 000160	+ 00.221682	+ 00.120107	+ 02	+ 03
- 003198	+ 000184	+ 00.222897	+ 00.161141	+ 02	+ 04
- 003434	+ 000125	+ 00.224559	+ 00.203036	+ 02	+ 05
- 003781	- 000133	+ 00.226759	+ 00.245997	+ 02	+ 06
- 004293	- 000821	+ 00.229617	+ 00.290185	+ 02	+ 07
- 005083	- 002439	+ 00.233290	+ 00.335592	+ 02	+ 08
- 003143	+ 000000	+ 00.261500	+ 00.000000	+ 03	+ 00
- 003169	+ 000010	+ 00.261702	+ 00.040140	+ 03	+ 01
- 003251	+ 000017	+ 00.262316	+ 00.080501	+ 03	+ 02
- 003400	+ 000007	+ 00.263368	+ 00.121307	+ 03	+ 03
- 003638	- 000050	+ 00.264902	+ 00.162789	+ 03	+ 04
- 004002	- 000226	+ 00.266983	+ 00.205180	+ 03	+ 05
- 004555	- 000659	+ 00.269707	+ 00.248710	+ 03	+ 06
- 005416	- 001639	+ 00.273197	+ 00.293564	+ 03	+ 07
- 003677	+ 000000	+ 00.304058	+ 00.000000	+ 04	+ 00
- 003716	- 000053	+ 00.304310	+ 00.040606	+ 04	+ 01
- 003840	- 000119	+ 00.305076	+ 00.081444	+ 04	+ 02

Portion of Biotar 12⁰ Predicted Displacements

$\epsilon'_y [2]$	$\epsilon'_z [2]$	$\epsilon'_y [3]$	$\epsilon'_z [3]$	y	z
-002666	+000000	-00.002678	+00.000000	+00	+00
-002676	+000108	-00.002689	+00.000109	+00	+01
-002708	+000222	-00.002720	+00.000223	+00	+02
-002760	+000344	-00.002770	+00.000339	+00	+03
-002830	+000468	-00.002836	+00.000439	+00	+04
-002916	+000577	-00.002915	+00.000483	+00	+05
-003014	+000638	-00.003000	+00.000389	+00	+06
-003120	+000598	-00.003081	+00.000023	+00	+07
-003230	+000383	-00.003144	-00.000830	+00	+08
-002721	+000000	-00.002723	+00.000000	+01	+00
-002735	+000083	-00.002737	+00.000085	+01	+01
-002778	+000173	-00.002780	+00.000175	+01	+02
-002850	+000272	-00.002854	+00.000266	+01	+03
-002951	+000374	-00.002962	+00.000341	+01	+04
-003081	+000463	-00.003108	+00.000359	+01	+05
-003240	+000506	-00.003299	+00.000241	+01	+06
-003429	+000451	-00.003541	-00.000148	+01	+07
-003648	+000225	-00.003842	-00.001020	+01	+08
-002875	+000000	-00.002858	+00.000000	+02	+00
-002893	+000051	-00.002877	+00.000053	+02	+01
-002948	+000108	-00.002935	+00.000108	+02	+02
-003039	+000174	-00.003038	+00.000160	+02	+03
-003171	+000242	-00.003195	+00.000188	+02	+04
-003345	+000296	-00.003420	+00.000149	+02	+05
-003565	+000302	-00.003729	-00.000040	+02	+06
-003836	+000209	-00.004147	-00.000517	+02	+07
-004164	-000058	-00.004701	-00.001497	+02	+08
-003179	+000000	-00.003146	+00.000000	+03	+00
-003201	+000011	-00.003172	+00.000009	+03	+01
-003268	+000027	-00.003254	+00.000015	+03	+02
-003382	+000050	-00.003399	+00.000009	+03	+03
-003546	+000071	-00.003625	-00.000036	+03	+04
-003768	+000074	-00.003952	-00.000169	+03	+05
-004053	+000024	-00.004411	-00.000480	+03	+06
-004411	-000131	-00.005043	-00.001113	+03	+07
-003688	+000000	-00.003684	+00.000000	+04	+00
-003715	-000038	-00.003722	-00.000054	+04	+01
-003796	-000072	-00.003840	-00.000118	+04	+02

APPENDIX 4

Formulae Used in Aberration Coefficients Programme

$$\begin{aligned}
y_p &= {}^*y_p - dv_p \\
i_p &= cy_p - v_p \\
i'_p &= ki_p \\
v'_p &= cy_p - i'_p \\
q &= i_q / i_p \\
j &= v_q - qv_p \\
\tilde{\omega} &= (k - 1)i_p j \\
1/N &= y_p v_q - y_q v_p \\
a_p &= \frac{1}{2}N(k - 1)y_p i_p^2 (v_p - i'_p) & A_p &= \sum_{i=1}^j a_p \\
\bar{a}_p &= qa_p & \bar{A}_p &= \sum_{i=1}^j \bar{a}_p \\
\bar{b}_p &= 2q\bar{a}_p & \bar{B}_p &= \sum_{i=1}^j \bar{b}_p \\
c_p &= \frac{1}{2}(\bar{b}_p - \tilde{\omega}) & &= \\
\bar{c}_p &= qc_p & &= \\
a_q &= \bar{a}_p - \frac{1}{2}(v_p'^2 - v_p^2) & &= \\
\bar{a}_q &= qa_q & &= \\
b_q &= 2\bar{a}_q + \tilde{\omega} & &= \\
c_q &= \frac{1}{2}qb_q & &= \\
\bar{c}_q &= qc_q & &= \\
A^+ &= qA_p - A_q \\
\bar{A}^+ &= q\bar{A}_p - \bar{A}_q \\
B^+ &= 2q\bar{A}_p - B_q \\
\bar{B}^+ &= q\bar{B}_p - 2C_q \\
C^+ &= qC_p - C_q \\
\bar{C}^+ &= q\bar{C}_p - \bar{C}_q \\
\mathcal{A} &= \bar{A}^+ - qA^+ \\
\mathcal{B} &= \bar{B}^+ - qB^+ \\
\mathcal{C} &= \bar{C}^+ - qC^+ \\
\underline{w}_p &= \frac{1}{8} [i_p^2 + i_p'^2 + v_p'^2 - 3v_p^2] \\
\underline{w}_q &= \frac{1}{8} [i_q^2 + i_q'^2 + v_q'^2 - 3v_q^2]
\end{aligned}$$

$$\begin{aligned}
\Phi_1 &= (k-1)v_p i_p \\
\Phi_2 &= \frac{1}{2}[8w_p + 3\Phi_1 - v_p^2 - ki_p^2] \\
\Phi_3 &= 2v_p^2 + 8w_p + 5\Phi_1 - \Phi_2 \\
\underline{s}_{1p} &= 3a_p w_p \\
s_{1p}^\dagger &= 2a_p A^\dagger + \underline{s}_{1p} \\
s_{1p} &= a_p A^\dagger + s_{1p}^\dagger & S_{1p} &= \sum_1^{j-1} s_{1p} \\
\bar{s}_{1p} &= a_p \mathcal{A} + q s_{1p} & &'' \\
w_1 &= (v_p^2 - 3v_p) j a_p \\
\underline{s}_{2p} &= 4q \underline{s}_{1p} + w_1 \\
s_{2p}^\dagger &= 2(qA^\dagger + \bar{A}^\dagger + B^\dagger)a_p - \bar{\omega} A_p + \underline{s}_{2p} \\
s_{2p} &= a_p (2qA^\dagger + B^\dagger) + s_{2p}^\dagger & S_{2p} &= \sum_1^{j-1} s_{2p} \\
\bar{s}_{2p} &= 2\bar{a}_p \mathcal{A} + a_p \mathcal{B} + q s_{2p} & &'' \\
w_2 &= \frac{1}{2} j^2 a_p \\
w_3 &= \frac{1}{4} (v_p^2 - v_p^2) v_p^2 j \\
\underline{s}_{3p} &= \frac{1}{2} q (w_1 + s_{2p}) + w_2 + w_3 \\
s_{3p}^\dagger &= 2(q\bar{A}^\dagger + C^\dagger)a_p - \bar{\omega} \bar{A}_p + \underline{s}_{3p} \\
s_{3p} &= c_p A^\dagger + a_p C^\dagger + s_{3p}^\dagger & S_{3p} &= \sum_1^{j-1} s_{3p} \\
\bar{s}_{3p} &= a_p C + c_p \mathcal{A} + q s_{3p} & &'' \\
\underline{s}_{4p} &= 2(\underline{s}_{3p} - 2w_2 - w_3) \\
s_{4p}^\dagger &= 2(s_{3p}^\dagger - \underline{s}_{3p}) + \underline{s}_{4p} \\
s_{4p} &= 2\bar{a}_p B^\dagger + s_{4p}^\dagger & S_{4p} &= \sum_1^{j-1} s_{4p} \\
\bar{s}_{4p} &= 2\bar{a}_p \mathcal{B} + q s_{4p} & &'' \\
w_4 &= \frac{1}{2} \bar{\omega}^2 \\
\underline{s}_{5p} &= q(2s_{3p} + \underline{s}_{4p} - q\underline{s}_{2p} - 2w_2) + w_4 \\
s_{5p}^\dagger &= 2[q(\bar{B}^\dagger + C^\dagger) + \bar{C}^\dagger]a_p + \underline{s}_{5p} - \bar{\omega}(\bar{B}_p + C_p) \\
s_{5p} &= 2\bar{a}_p C^\dagger + c_p B^\dagger + s_{5p}^\dagger & S_{5p} &= \sum_1^{j-1} s_{5p} \\
\bar{s}_{5p} &= 2a_p C + c_p \mathcal{B} + q s_{5p} & &'' \\
w_5 &= \frac{1}{8} j^2 \bar{\omega} \\
\underline{s}_{6p} &= \frac{1}{2} [\frac{1}{4} q (2w_3 - 2\underline{s}_{3p} - \underline{s}_{4p}) + w_4 + \underline{s}_{5p}] q + w_5 \\
s_{6p}^\dagger &= 2[c_p \bar{C}_p - \bar{a}_p \bar{C}_q] + \underline{s}_{6p} \\
s_{6p} &= c_p C^\dagger + s_{6p}^\dagger & S_{6p} &= \sum_1^{j-1} s_{6p} \\
\bar{s}_{6p} &= c_p C + q s_{6p} & &''
\end{aligned}$$

$$\begin{aligned}
S_{1q} &= \bar{A}_p (v_p^2 - A_q) - \frac{3}{2} (\bar{A}_p - A_q)^2 + A_p (\bar{A}_q - v_p v_q) + \bar{S}_{1p} \\
S_{2q} &= \bar{A}_p (2\bar{A}_q - B_q - v_p v_q) + \bar{A}_q v_p^2 - A_q (v_p v_q + \bar{B}_p) + A_p (2C_q - v_q^2) \\
&\quad + \bar{B}_p v_p^2 - 3(\bar{B}_p - B_q)(\bar{A}_p - A_q) + \bar{S}_{2p} \\
S_{3q} &= C_p \bar{A}_q + A_p \bar{C}_q - \bar{A}_p C_q - A_q (\bar{C}_p + v_q^2) - \frac{1}{2} v_p v_q (\bar{B}_p - B_q) + v_p^2 \bar{C}_p \\
&\quad - 3(\bar{A}_p - A_q)(\bar{C}_p - C_q) + \bar{S}_{3p} \\
S_{4q} &= 2[\bar{A}_p (2C_q - v_q^2) + v_p^2 C_q] - B_q \bar{B}_p - (\bar{B}_p - B_q) [\frac{3}{2} (\bar{B}_p - B_q) - v_p v_q] + \bar{S}_{4p} \\
S_{5q} &= C_q (2C_p - \bar{B}_p + v_p v_q) + \bar{C}_p v_p v_q - B_q (\bar{C}_p + v_q^2) + \bar{C}_q (2\bar{A}_p + v_p^2) \\
&\quad + \bar{S}_{5p} - 3(\bar{B}_p - B_q)(\bar{C}_p - C_q) - C_p v_q^2 \\
S_{6q} &= \bar{C}_q (C_p + v_p v_q) - C_q (\bar{C}_p + v_q^2) - \frac{3}{2} (\bar{C}_p - C_q)^2 + \bar{S}_{6p} \\
\bar{S}_{1q} &= \frac{1}{4} [2S_{3p} + 6(A_p C_q - \bar{A}_p C_p + \bar{A}_q [\bar{A}_p - A_q]) - v_q (A_p v_q - A_q v_p) \\
&\quad + v_p (\bar{A}_p v_q - \bar{A}_q v_p)] \\
\bar{S}_{2q} &= S_{5p} + 4\bar{A}_p C_q - \bar{A}_q \bar{B}_p - B_q C_p - (\bar{B}_p - B_q)^2 \\
&\quad + 2[C_q (2\bar{A}_p - A_q) - \bar{A}_q C_p + A_p \bar{C}_q - \bar{A}_p \bar{C}_p] \\
\bar{S}_{3q} &= \frac{1}{2} [4 S_{6p} + 6 (\bar{A}_p \bar{C}_q - C_q \bar{A}_q - C_p [\bar{C}_p - C_q]) + v_q (C_p v_q - C_q v_p) \\
&\quad - v_p (\bar{C}_p v_q - \bar{C}_q v_p)] \\
\bar{S}_{4q} &= S_{5q} + 2(\bar{B}_p C_q - 2C_q C_p) + (2\bar{A}_p \bar{C}_q - B_q \bar{C}_p) + v_p (C_q v_q - \bar{C}_q v_p) \\
&\quad + \frac{1}{2} v_q (B_q v_q - 2C_q v_p) \\
\bar{S}_{5q} &= 2 [2S_{6q} + \bar{C}_p C_q - \bar{C}_q C_p + v_q (C_q v_q - \bar{C}_q v_p)] + 3(\bar{B}_p \bar{C}_q - 2C_q \bar{C}_p) \\
\bar{S}_{6q}^* &= \frac{1}{2} 3 C_q [q (\bar{C}_p + \underline{w}_q) - \bar{C}_q] \quad * = \text{at next surface} \\
&\quad \text{i.e. } (j+1)^{\text{th}}
\end{aligned}$$

$$S_1^+ = q S_{1p} - S_{1q}$$

$$\bar{S}_1^+ = q \bar{S}_{1p} - \bar{S}_{1q}$$

$$S_2^+ = q S_{2p} - S_{2q}$$

$$\bar{S}_2^+ = q \bar{S}_{2p} - \bar{S}_{2q}$$

"

"

"

$$\bar{S}_6^+ = q \bar{S}_{6p} - \bar{S}_{6q}$$

$$\mathcal{S}_1 = \bar{S}_1^+ - qS_1^+$$

$$\mathcal{S}_2 = \bar{S}_2^+ - qS_2^+$$

"

"

"

$$\mathcal{S}_6 = \bar{S}_6^+ - qS_6^+$$

$$z_1 = a_p [10\underline{w}_p^2 + \frac{1}{4}ja_p(2v_p^+ - 5v_p)] = \underline{t}_{1p}$$

$$z_2 = \frac{1}{2} [10\underline{w}_p w_1 - ja_p v_p (i_p + \underline{v})(i_p^+ - v_p)]$$

$$z_3 = -\frac{3}{16}\bar{\omega} [4\underline{w}_p \Phi_3 + \Phi_1 (3\Phi_1 - 2\Phi_2 - 8\underline{w}_p)]$$

$$z_4 = \frac{1}{4} w_2 (2\Phi_2 - 10\underline{w}_p + 5v_p^2)$$

$$z_5 = -\frac{1}{2} \bar{\omega} j \Phi_2 (v_p + v_p^+)$$

$$z_6 = 2w_5 (\underline{w}_p + v_p^2 - \frac{5}{4} \Phi_1)$$

$$z_7 = -j^2 w_1$$

$$z_8 = -2w_5 (5\Phi_1 + 4\underline{w}_p - 2v_p^2)$$

$$z_9 = -\frac{3}{4} j^2 w_4$$

$$z_{10} = -\frac{1}{2} j^2 w_5$$

$$t_{1p} = 4 [\frac{1}{8} A_p (2s_{2p} - \bar{\omega} A_p) - A_q \underline{s}_{1p}] + a_p (A^{\dagger^2} + 3S_1^+) + A^+ s_{1p}^+ + z_1$$

$$\bar{t}_{1p} = q t_{1p} + a_p \mathcal{S}_1 + A s_{1p}^+$$

$$t_{2p} = \underbrace{(6qz_1 + z_2)}_{t_{2p}} + 4a_p [\frac{1}{2} \{ A^+ (\bar{A}^+ + B^+) + \bar{S}_1^+ + \frac{3}{2} S_2^+ \} + qS_1^+] \\ + 3 [\underline{s}_{2p} (\bar{A}_p - A_q) - \bar{\omega} (A_p \bar{A}_p + \frac{1}{3} S_{1p})] + 2A_p (\underline{s}_{3p} + \underline{s}_{4p}) + A^+ s_{2p}^+ \\ + B^+ s_{1p}^+ - 4\underline{s}_{1p} (\bar{A}_q + B_q)$$

$$\bar{t}_{2p} = q(2a_p \mathcal{S}_1 + t_{2p}) + a_p \mathcal{S}_2 + A s_{2p}^+ + B s_{1p}^+$$

$$t_{3p} = \underbrace{\{ \frac{1}{2} q [\underline{t}_{2p} + z_2] + z_3 + z_4 \}}_{t_{3p}} + a_p [2(q\bar{S}_1^+ + A^+ C^+) + \bar{A}^{\dagger^2} + 3S_3^+] \\ - \bar{\omega} (\frac{1}{2} \bar{A}_p^2 + A_p C_p + \bar{S}_{1p}) + A^+ s_{3p}^+ + C^+ s_{1p}^+ + 2 [\underline{s}_{3p} (\bar{A}_p - A_q) - 2C_q \underline{s}_{1p}] \\ + A_p \underline{s}_{3p} + \frac{1}{2} \underline{s}_{2p} (\bar{B}_p - B_q) + c_p S_1^+$$

$$\bar{t}_{3p} = q t_{3p} + a_p \mathcal{S}_3 + c_p \mathcal{S}_1 + A s_{3p}^+ + C s_{1p}^+$$

$$t_{4p} = \underbrace{\{ 4(\underline{t}_{3p} - z_3 + z_4) \}}_{t_{4p}} + a_p [2(2qS_2^+ + A^+ \bar{B}^+ + \bar{A}^+ B^+) + B^{\dagger^2} + 2\bar{S}_2^+ + 3S_4^+] \\ - \bar{\omega} (4\bar{A}_p^2 + A_p \bar{B}_p + S_{2p}) + 4(\bar{A}_p \underline{s}_{3p} - 2C_q \underline{s}_{1p}) + B^+ s_{2p}^+ + A^+ s_{4p}^+ \\ + 2 [\underline{s}_{4p} (3\bar{A}_p - A_q) + A_p \underline{s}_{3p}] + \underline{s}_{2p} (\bar{B}_p - 2\bar{A}_q - 3B_q)$$

$$\bar{t}_{4p} = a_p (2q \mathcal{S}_2 + \mathcal{S}_4) + q t_{4p} + A s_{4p}^+ + B s_{2p}^+$$

$$\begin{aligned}
t_{5p} &= \overbrace{\{q(t_{4p} + 4t_{3p} + 8z_4 - 2qt_{2p}) + z_5\}}^{t_{5p}} + 2a_p [q(2S_3^+ + \bar{S}_2^+) + \bar{S}_3^+ + \frac{3}{2}S_6^+] \\
&\quad + S_2^+ c_p + 2a_p [\bar{A}^+ (\bar{B}^+ + C^+) + A^+ \bar{C}^+ + B^+ C^+] - \bar{\omega} [\bar{A}_p (3C_p + \bar{B}_p) + A_p \bar{C}_p + \bar{S}_2 p + S_{3p}] \\
&\quad + \underline{S}_{2p} (\bar{C}_p - 5C_q) + C^+ s_{2p}^+ (\bar{B}_p - B_q) (3\underline{S}_{3p} + \underline{S}_{4p}) + B^+ s_{3p}^+ + \underline{S}_{5p} (5\bar{A}_p - A_q) \\
&\quad + A^+ s_{5p}^+ + 4(A_p \underline{S}_{6p} - \bar{C}_q \underline{S}_{1p}) \\
\bar{t}_{5p} &= A s_{5p}^+ + B s_{3p}^+ + C s_{2p}^+ + a_p (2q\delta_3 + \delta_5) + q t_{5p} + c_p \delta_2 \\
t_{6p} &= \overbrace{\{q [q(z_3 + z_4 - t_{3p}) + \frac{1}{2}t_{5p} + \frac{1}{2}z_5] + z_6\}}^{t_{6p}} + a_p [2(q\bar{S}_3^+ + \bar{A}^+ \bar{C}^+) + C^{+2} + 3S_6^+] \\
&\quad - \bar{\omega} (\frac{1}{2}C_p^2 + \bar{A}_p \bar{C}_p + \bar{S}_3 p) + C^+ s_{3p}^+ + A^+ s_{6p}^+ + 2(2\bar{A}_p \underline{S}_{6p} + \underline{S}_{3p} (\bar{C}_p - C_q) \\
&\quad - \bar{C}_q \underline{S}_{2p} + c_p S_3^+ + \frac{1}{2}(\bar{B}_p - B_q) \underline{S}_{5p} \\
\bar{t}_{6p} &= q t_{6p} + a_p \delta_6 + c_p \delta_3 + A s_{6p}^+ + C s_{3p}^+ \\
t_{7p} &= \overbrace{\{\frac{2}{3}[4q(z_4 - z_3) + t_{5p} - z_5] + z_7\}}^{t_{7p}} + 2a_p (2q S_4^+ + B^+ \bar{B}^+ + \bar{S}_4^+) \\
&\quad + 2\underline{S}_{4p} (\bar{B}_p - B_q) - \bar{\omega} (2\bar{A}_p \bar{B}_p + S_{4p}) + B^+ s_{4p}^+ + 4(\bar{A}_p \underline{S}_{5p} - C_q \underline{S}_{2p}) \\
\bar{t}_{7p} &= q(2a_p \delta_4 + t_{7p}) + B s_{4p}^+ \\
t_{8p} &= \overbrace{\{q[4q(z_4 - z_3) + 3z_7 - 2z_5] + z_8 + 4(t_{6p} - z_6)\}}^{t_{8p}} \\
&\quad - \bar{\omega} [(\frac{1}{2}\bar{B}_p + C_p) \bar{B}_p + \bar{S}_{4p} + S_{5p}] + c_p S_4^+ + 2\bar{C}_p (\underline{S}_{4p} - \bar{\omega} \bar{A}_p) + C^+ s_{4p}^+ \\
&\quad + B^+ s_{5p}^+ + a_p [2\{q(2S_5^+ + \bar{S}_4^+) + \bar{B}^+ C^+ + B^+ \bar{C}^+ + \bar{S}_5^+\} + \bar{B}^{+2}] \\
&\quad + 8[\bar{A}_p \underline{S}_{6p} - \frac{1}{4}\{C_q (2\underline{S}_{3p} + 3\underline{S}_{4p}) + \bar{C}_q \underline{S}_{2p}\}] + \underline{S}_{5p} (3\bar{B}_p + 2C_p - B_q) \\
\bar{t}_{8p} &= q(2a_p \delta_5 + t_{8p}) + c_p \delta_4 + B s_{5p}^+ + C s_{4p}^+ \\
t_{9p} &= q[q(4qz_4 + z_5 + \frac{3}{4}(z_7 - t_{7p}) + t_{8p} + 2z_6 + z_8) + z_9 \\
&\quad + 2a_p [q(2S_6^+ + \bar{S}_5^+) + \bar{C}^+ (\bar{B}^+ + C^+) + \bar{S}_6^+] + c_p S_5^+ + C^+ s_{5p}^+ \\
&\quad + B^+ s_{6p}^+ - \bar{\omega} [\bar{C}_p (\bar{B}_p + C_p) + \bar{S}_{5p} + S_{6p}] + 4[\underline{S}_{6p} (\bar{B}_p + C_p) + \frac{3}{4} \underline{S}_{5p} (\bar{C}_p - C_q)] \\
&\quad - 2\bar{C}_q (\underline{S}_{3p} + \underline{S}_{4p}) \\
\bar{t}_{9p} &= q(2a_p \delta_6 + t_{9p}) + c_p \delta_5 + B s_{6p}^+ + C s_{5p}^+ \\
t_{10p} &= q[q\{q[q(qz_4 + z_5) + z_6 + 9z_4] + z_7 + z_8\} + z_9] + z_{10} \\
&\quad + a_p (2q\bar{S}_6^+ + \bar{C}^{+2}) + c_p S_6^+ - \bar{\omega} (\frac{1}{2}\bar{C}_p^2 + \bar{S}_{6p}) + C^+ s_{6p}^+ \\
&\quad - \bar{C}_q \underline{S}_{5p} + 4\bar{C}_p \underline{S}_{6p} \\
\bar{t}_{10p} &= q t_{10p} + c_p \delta_6 + C s_{6p}^+
\end{aligned}$$

Final tertiary aberration coefficients

$$\begin{aligned}
T'_{ipk} &= \sum_i^k t_{ip} \\
\bar{T}'_{ipk} &= \sum_i^k \bar{t}_{ip}, \text{ etc.}
\end{aligned}$$

REFERENCES

1. H. A. Buchdahl, Optical Aberration Coefficients (Oxford University Press, London 1954).
2. H. A. Buchdahl, J. Opt. Soc. Am. 46, 941 (1956).
3. H. A. Buchdahl, J. Opt. Soc. Am. 48, 563 (1958).
4. H. A. Buchdahl, J. Opt. Soc. Am. 48, 747 (1958).
5. H. A. Buchdahl, J. Opt. Soc. Am. 48, 757 (1958).
6. H. A. Buchdahl, J. Opt. Soc. Am. 49, 1113 (1959).
7. H. A. Buchdahl, J. Opt. Soc. Am. 50, 534 (1960).
8. H. A. Buchdahl, J. Opt. Soc. Am. 50, 539 (1960).
9. H. A. Buchdahl, J. Opt. Soc. Am. 50, 678 (1960).
10. F. D. Cruickshank and G. A. Hills, J. Opt. Soc. Am. 50, 379 (1960).
11. M. Herzberger, J. Opt. Soc. Am. 37, 485 (1947).
12. Y. Ukita and J. Tsujiuchi, Optica Acta, 5, 39 (1958).

PUBLISHED WORK

"Use of a Digital Computer for the Calculation of Aberration Coefficients."

J. Opt. Soc. Am. 49, 875 (1959).

"New Ray Tracing Scheme,"

J. Opt. Soc. Am. 50, 528 (1960).

"Optical Aberration Coefficients and Optical Design,"

Proceedings of Conference on Optical Instruments and Techniques, London, 1961.

Use of a Digital Computer for the Calculation of Aberration Coefficients

P. W. FORD

Physics Department, University of Tasmania, Hobart, Australia

(Received February 2, 1959)

A scheme has been published which enables the computation of the aberration coefficients of axially symmetrical optical systems, up to and including the seventh order, using a desk calculating machine. The present paper describes the design and performance of a program for a particular digital computer based on this scheme.

1. INTRODUCTION

THE successful design of axially symmetrical optical systems in this laboratory¹ results from the use of initial design methods developed by Cruickshank,² together with the final balancing of aberrations by the consideration of higher order aberration coefficients; a theory of these which enables them to be computed having been developed by Buchdahl³ and extended in recent papers.⁴⁻⁷ In III, a scheme is presented which enables one to compute all the coefficients up to and including the tertiary (seventh order) coefficients (Table I). In IV this table is extended to give the quaternary spherical aberration coefficient. Although Table I of III is expressly designed for a particular type of desk computer, it nevertheless involves a considerable amount of tedious work. In order to make the coefficients more readily available for use in design work, a serial digital computer, situated at the University of New South Wales, Sydney, has been programed to compute all the coefficients for which Table I of III was designed. This computer, described in Sec. 2, has proved quite satisfactory, possessing adequate storage and high speed operation. The time taken to compute and punch out all the coefficients (41 in number) for one surface of a system is 45 sec, which compares favorably with other less comprehensive work in this field.⁸

2. COMPUTER DESCRIPTION

The computer is a standard English Electric Deuce machine. It has a high speed store of 402 words and a magnetic drum store of 8192 words. The high speed store consists of mercury lines of various lengths arranged to minimize access time. Transfers between the drum and the high speed store occur in blocks of 32 words, each of these transfers occupying one drum revolution (about 13 msec). The read head for the drum

consists of a bank of 16 heads which can be moved to any one of 16 positions along the drum, about 50 msec being required for the shift. The writing head is an identical unit on the opposite side of the drum. The transfers and head shifts take place without affecting the normal operation of the rest of the machine and are generally arranged so as to cause no delay in access to the drum.

The arithmetic facilities are distributed over four immediate access (single word) registers and one double length unit. This enables some partial results to be retained in the arithmetic section. Multiplication and division, associated with a single and a double length store, proceed at the rate of almost 30 000 per minute and continue automatically once they are setup. Thus the machine can, at the same time, be carrying out logical operations, shifting, adding, subtracting, or discriminating. Additions and subtractions proceed at the rate of almost 1 000 000 per min, with about twice this speed for the special case of summing data stored sequentially.

The program and data are punched on standard 80-column Hollerith cards, which are read in at the rate of 200 cards per min, and punched out at the rate of 100 cards per min. The digit frequency is 1 Mc/sec and one word contains 32 digits. The machine uses a modified 2-address system; an instruction indicating the source and destination of a transfer, the duration of the transfer and the location of the next instruction in the high speed store.

3. COMPUTING SCHEME USED FOR PROGRAMING

The scheme used for the program is the condensed computing scheme for systems containing spherical surfaces only, given in Table I of III. This scheme gives the first, third, fifth, and seventh order a -coefficients for such a system. The a -coefficients alone will define the displacement, if, and only if the paracanonical coordinates appropriate to the given positions of the object and entrance pupil planes are used, M, Sec. 12(a). Thus, if as is the case here, one wishes to calculate the seventh order aberrations, the third and fifth order b -coefficients are calculated only because they are required in obtaining the seventh order a -coefficients.*

* As is mentioned later, the condition that N_1/g be unity is, in practice, always imposed as a matter of convenience. In that case, according to M, Sec. 12(b), there is no formal distinction between

¹ Cruickshank and Hills (in preparation).

² F. D. Cruickshank, *Australian J. Phys.* 11, 41 (1958).

³ H. A. Buchdahl, *Optical Aberration Coefficients* (Oxford University Press, New York, 1954), hereafter to be referred to as M.

⁴ H. A. Buchdahl, *J. Opt. Soc. Am.* 46, 941 (1956), hereafter to be referred to as I.

⁵ H. A. Buchdahl, *J. Opt. Soc. Am.* 48, 563 (1958), hereafter to be referred to as II.

⁶ H. A. Buchdahl, *J. Opt. Soc. Am.* 48, 747 (1958), hereafter to be referred to as III.

⁷ H. A. Buchdahl, *J. Opt. Soc. Am.* 48, 757 (1958), hereafter to be referred to as IV.

⁸ E. g., D. P. Feder, *J. Opt. Soc. Am.* 41, 630 (1951).

The basic equations given in M, Sec. 77–81 for the tertiary coefficients, while directly suitable for systems containing aspherical surfaces, require a storage capacity in excess of that available in Deuce. Certain groups of terms occur more than once in these equations, and, in order to use these groups to avoid unnecessary calculation, quick access storage for approximately 300 words is required. Including the tertiary results, this exceeds the storage capacity since two of the mercury lines contain “permanent” instructions used throughout the entire program and one line must be kept free for drum transfers. If none of the repeated groups are used except the 55 products (M, page 145), the storage can be arranged, but approximately 1295 arithmetical operations are required. Compared with this, storage for 45 terms and only 524 operations are needed using the condensed scheme. In view of this, and considering that systems containing aspherical surfaces are comparatively rare, the condensed scheme is by far the most practical method for calculating the aberration coefficients.

Nevertheless, it was found desirable, from the point of view of storage, to deviate occasionally from the scheme as presented in Table I of III. In the program, the entries t_{101} to t_{114} of Table I are calculated from the definitive equations M (84.15)—namely,

$$S_{\mu}^{\dagger} = qS_{\mu p} - S_{\mu q},$$

$$\bar{S}_{\mu}^{\dagger} = q\bar{S}_{\mu p} - \bar{S}_{\mu q},$$

where the S_{μ}^{\dagger} are the $S_{\mu(T)}$ in M, Sec. 84.

In order to use these equations, the $\bar{S}_{\mu q}$ are required, and these are not calculated in Table I of III. These coefficients have been obtained from the identities given in M, pages 32, 33, except for \bar{s}_{6q} , which is obtained from the last member of the Eqs. M (11.3). This equation can be rewritten as

$$\bar{s}_{6q} = 3c_q [q\{\bar{C}_p + \frac{1}{8}(i_q^2 + i_q'^2 + v_q'^2 - 3v_q^2)\} - \bar{C}_q].$$

The term $\frac{1}{8}(i_q^2 + i_q'^2 + v_q'^2 - 3v_q^2)$ has the same form as w in Table I of III and hence is calculated by the same part of the program that calculates w , but using the q quantities as initial data, the result being stored until the \bar{s}_{6q} is calculated. Since \bar{s}_{6q} requires c_q , the $A_q \cdots \bar{C}_q$ are obtained from the $a_q \cdots \bar{c}_q$, the latter being obtained from the Eqs. M (20.21), (20.33), (20.34). This also avoids retaining the t_6 of the previous surface used in the expression for \bar{C}_q in Table I of III.

In the condensed scheme, the condition that N_1/g be unity is imposed as a matter of convenience, since all the intermediate coefficients are divided by N_1/g , (M, page 17.) Using $\bar{O}T$ coordinates, (M, page 18) g is necessarily unity, and hence it is only necessary to ensure that N_1 is also unity. Since most systems work in air, this requires no further manipulation, but should N_1 not be unity, then all the refractive indices which occur

all the expressions and equations relating to canonical coefficients (distinguished by subscripts p and q) on the one hand, and to paracanonical coefficients (distinguished by subscripts a and b) on the other. Following precedent, p and q will therefore be used throughout except where reference is made to initial input data.

in the system must be multiplied by a common factor so chosen to make N_1 unity. In order to get the actual coefficients describing the displacement, it is necessary to multiply the final coefficients by $\mu (= 1/N_k' v_{pk}')$. Also for convenience, the contributions to these by the surfaces of the system are multiplied by μ (M, Sec. 7).

If a system is working in air with an object at infinity and unit focal length, then $\mu = 1$. For such a system, the actual coefficients describing the image defects and the contributions to these by the surfaces of the system *for all rays* are just those as computed by Table I of III. Since this may not always be the case, i.e., $\mu \neq 1$, the coefficients are all multiplied, in the Deuce program, by μ before being punched out.

The input data for the program are (i) curvatures, $c_j (j=1 \cdots k)$; (ii) separations, $d_j (d_1=0)$; (iii) ratio of refractive indices, $k (= N_j/N_j')$; (iv) a and b ray data, i.e., $v_{a1}, y_{a1}, v_{b1}, y_{b1}$; (v) N_k' .

Using $\bar{O}T$ coordinates, and having an object at infinity, $v_{a1}=0, y_{a1}=1, v_{b1}=1, y_{b1}=p$, where p is the distance of the entrance pupil from the first tangent plane [M, Sec. 33(a)]. With the insertion of these data, in floating decimal, in correct order (as presented above), the machine proceeds through the system, one surface at a time. At each surface, all the augmented coefficients up to \bar{i}_{10pj} and the sums of the coefficients of the kind A_{pj}' to \bar{T}_{10pj}' are calculated, before going to the next surface. At the last surface, the $A_{pk}' \cdots \bar{T}_{10pk}'$ are the augmented aberration coefficients of the entire system, (M, Sec. 8). Inserted in binary as part of the program are the coefficients of the second surface of the system computed in Table I of III. When the machine determines that an entire system has been computed, it proceeds to a check program which computes the coefficients of the surface already mentioned. It then compares the two sets of coefficients in groups $(\bar{i}, \bar{i}', A, S, T, \bar{T})$ and, if they agree, proceeds to the punch out routine. If they do not agree, the failure is indicated on the panel of the machine by a combination of lights which locates the group in which the error has occurred. The input data are also checked to see that they have all been correctly used. These checks are designed to guard against any consistent machine failure which could occur during the main program, such as an arithmetic failure. Random errors, such as the insertion or loss of digits in the results cannot be adequately checked in this fashion, but since faults of this kind usually also affect the program instructions, they become self evident.

The following results are punched out in floating decimal, *for each surface of the system*: (i) paraxial coefficients, $v_{pj}', y_{pj}, v_{qj}', y_{qj}$, (first order); (ii) primary coefficients, (third order); (iii) secondary coefficients, (fifth order); (iv) tertiary coefficients, (seventh order); and finally (v) the third, fifth, seventh order coefficients for the entire system. The coefficients (ii), (iii), (iv), (v) are first multiplied by μ before being punched out. To assist in the layout of the typed results, “blank” cards

are inserted by the machine between each of the groups of coefficients.

The read-in time of the initial pack of program and data cards is approximately two minutes, and for subsequent systems, the "restore control" pack and data cards about fifteen seconds. The actual computing time per surface is eleven seconds, all arithmetic operations being performed in floating binary which enables the machine to handle numbers in the range $\pm 10 \exp(\pm 10^{10})$.

4. ACKNOWLEDGMENTS

I wish to express my thanks to the Electrical Engineering Department of the University of New South Wales, in particular to Mr. R. Smart and his staff associated with the computer, without whose generous assistance this program would not have been done. I also wish to thank Dr. H. A. Buchdahl for his assistance in the design of the program and Mrs. B. Brown who checked the coding and the cards.

New Ray Tracing Scheme

P. W. FORD

Department of Physics, University of Tasmania, Hobart, Tasmania, Australia

(Received December 1, 1959)

A detailed theoretical treatment is given of a new algebraic ray tracing scheme for tracing rays from an object point through any axially symmetrical optical system, including catadioptric systems, which contain spherical surfaces only. An explanation is given of the coordinate systems used which effect considerable simplifications both in the theory and its application. Also presented is a fully worked example of a skew ray trace through a wide angle system. An electronic computer has been programed for this ray trace and details of this are given. A predetermined scan interval is given to the machine which then proceeds automatically to trace from an object point rays spaced at this interval over the first polar tangent plane. For pencils of all obliquities, vignetting is carried out by the machine, which produces as one of the results the apparent shape of the entrance pupil.

1. INTRODUCTION

A NEW algebraic ray tracing scheme has been designed for tracing rays through any axially symmetrical optical system which contains spherical surfaces only. The scheme was designed by Buchdahl on the basis of certain equations contained in his monograph *Optical Aberration Coefficients*,¹ and hence the results from the ray trace can be directly compared with those predicted by these aberration coefficients. The design of the scheme is such that it gives some insight into the effect of each surface of a system on the final image.

For systems of unit focal length, the magnitude of the numbers appearing during the course of the ray trace are such that an electronic computer whose word length is ten decimal digits can work in fixed-point arithmetic to an accuracy of six decimal places. There are no trigonometric functions to be evaluated, and thus the scheme is quite suitable for fast electronic computation. It has been programed for a Deuce machine (described elsewhere²), the time taken to trace a ray through one surface being 0.5 sec.

Most of the symbols used in this paper have been defined in M, thus only those which do not appear in it will be defined here. The theory is presented in considerable detail to aid in the interpretation of M.

Consider a ray from an object point incident at the first polar tangent plane of an optical system. Then the ray whose subsequent path is calculated not by means of the true laws of refraction but instead by their linear approximations, i.e., the laws of paraxial optics, will be known as the "ideal ray." Now, the path of the actual ray, after refraction by the first surface, will not, in general, be identical with that of the ideal ray. Thus the difference between the coordinates of the intersection points of the actual ray and the ideal ray in the j th polar tangent plane will be a measure of the aberrations produced by the preceding ($j-1$) surfaces. Thus, if for

a particular ray the aberrations of the preceding surfaces are known, the coordinates of the intersection point of the ray with the polar tangent plane can be found from those of the ideal ray. This is stated in Eqs. M(9.3) where the δ 's, M(9.4), are the sum of the aberrations produced by the preceding surfaces. In Eq. M(9.4) $\Delta\Lambda_j$ is a measure of the aberrations produced by the j th surface alone, and Eq. M(7.5) shows that these can be summed over all the surfaces of a system to give, in the ideal (paraxial) image plane, the differences between the coordinates of the intersection point of the ray and those of the ideal image point. An ideal image point is defined as the intersection point of an ideal ray in the appropriate image plane.

Now it is not feasible to obtain an expression for $\Delta\Lambda_j$ in closed form, regarded as a function of the coordinates of a ray. In M, $\Delta\Lambda_j$ is therefore expressed as a pair of power series in the coordinates Y_1, V_1 , the coefficients of these series being called the (contributions to the) aberration coefficients, and these therefore characterize systems of rays rather than individual rays. For an individual ray one may of course calculate $\Delta\Lambda_1, \Delta\Lambda_2, \Delta\Lambda_3, \dots$ in turn by ordinary algebraic means, which amounts to tracing the ray through the system on the basis of Eqs. M(9.3) [or their equivalent M(12.9) in the case of paracanonical coordinates]. The $\Delta\Lambda$'s then being known, the value of ϵ_k' may be found immediately from Eq. M(7.5) which can be compared if desired with that computed from the aberration coefficients of M.

If paracanonical coordinates are used to specify a ray incident at the first polar tangent plane, then only the "a coefficients" occur in the actual description of the aberrations associated with the conjugate image planes (see M Sec. 12). This effects a considerable reduction in the number of calculations required to determine the aberrations of a system. Also, with the use of these coordinates, the form of the expressions derived in the ray trace theory is not dependent on the position of the object and entrance pupil planes. This makes it possible to design a simple "scan" procedure for the programed ray trace so that the machine will automatically produce

¹H. A. Buchdahl, *Optical Aberration Coefficients* (Oxford University Press, New York, 1954). This work will be referred to hereafter as M.

²P. W. Ford, J. Opt. Soc. Am. 49, 875 (1959).

the results for a "spot diagram" if desired. In view of the importance of the paracanonical coordinates in simplifying the ray trace computation as well as those for the aberration coefficients, the next section of this paper has been devoted to a discussion of these coordinate systems.

2. COORDINATE SYSTEMS

In M Sec. 5(a) it is shown that there exist relations between the paraxial variables of the type

$$\begin{aligned} Y_j &= y_{pj}Y_1 + y_{aj}V_1 \\ V_j &= v_{pj}Y_1 + v_{aj}V_1 \end{aligned}$$

where the coefficients are called "canonical paraxial coefficients," canonical coordinates (Fig. 1) being defined in M Sec. 3. For any optical system, the paraxial coefficients can be found by tracing two tangential paraxial rays whose initial data are; p ray, $y_1=1$, $v_1=0$; q ray, $y_1=0$, $v_1=1$. Then the values of y_j , v_j appearing in the p trace are y_{pj} , v_{pj} and those in the q trace y_{aj} , v_{aj} . One form of the paraxial relations used for the trace is

$$\begin{aligned} y_j &= {}^*y_j - d_j v_j \\ i_j &= c_j y_j - v_j \\ v_j' &= (1 - k_j) i_j + v_j \end{aligned}$$

where ${}^*y_j = y_{j-1}$, $v_j' = v_{j+1}$ and $k_j = N_j/N_{j+1}$. Note that by definition $d_1 = 0$.

Other paraxial variables can be expressed in terms of the paraxial coefficients and the ray parameters at the first surface. For example, consider an optical system whose entrance pupil is distant p from A , the pole of the first surface [see M, Sec. 33(a)]. Suppose the axial point of the object is l_{o1} from A , then from Eqs. M(5.101), (5.102)

$$\begin{aligned} h_j &= (N_1 v_{o1}/N_j v_{oj}) h_1 \\ &= N_1 v_1 (l_{o1} - p)/N_j (v_{pj} l_{o1} + v_{aj}) \end{aligned}$$

where v_1 refers to the principal ray.

Now if $p=0$ and the object is infinitely distant then $h_j = N_1 v_1/N_j v_{oj}$.

Similarly, many other expressions simplify when $p=0$ and $l_{o1} \rightarrow -\infty$.

It is possible, however, to effect the same simplifications no matter what the values of p and l_{o1} by choosing new coordinates systems in which to define the ray. The new coordinates of the ray are called "paracanonical coordinates."

If one chooses paracanonical coordinates S_j , S_n , T_j , T_n such that

$$\begin{aligned} S &= \sigma Y_1 + \bar{\sigma} V_1 \\ T &= \tau Y_1 + \bar{\tau} V_1 \end{aligned} \quad \text{M(12.1)}$$

where σ , $\bar{\sigma}$, τ , $\bar{\tau}$ are constants yet to be chosen, and $\bar{g} = \sigma \bar{\tau} - \bar{\sigma} \tau$, (\bar{g} is required to be nonzero) then there still

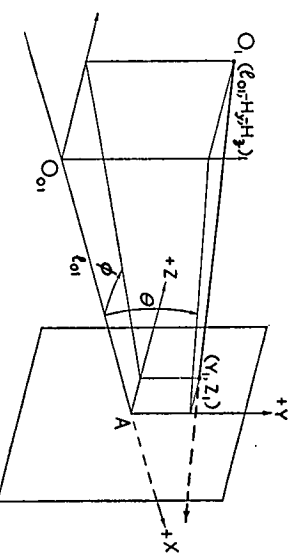


FIG. 1. Canonical coordinates. $V = \tan \theta$, $W = \tan \phi$.

exists relations between any paraxial linear variable

$$u_j = u_{aj}S + u_{bj}T \quad \text{M(12.3)}$$

where the u_{aj} , u_{bj} are called "paracanonical paraxial coefficients." The relations between these and the canonical coefficients are given by M(12.61, 62).

Now, since

$$\begin{aligned} Y_j &= y_{aj}S + y_{bj}T \\ V_j &= v_{aj}S + v_{bj}T \end{aligned}$$

from M(12.3), it is possible to calculate these coefficients in the same way as the canonical coefficients, only this time for the a ray, $s=1$, $l=0$; b ray, $s=0$, $l=1$. Hence, from M(12.1) the initial data for the paraxial traces are

$$\begin{aligned} a \text{ ray } y_1 &= \bar{\tau}/\bar{g}, & v_1 &= -\tau/\bar{g} \\ b \text{ ray } y_1 &= -\bar{\sigma}/\bar{g}, & v_1 &= \sigma/\bar{g} \end{aligned} \quad \text{M(12.5)}$$

and the values of y_j , v_j at each surface in the a ray trace will be y_{aj} , v_{aj} and those in the b ray trace y_{bj} , v_{bj} .

A particular set of paracanonical coordinates, called $\bar{O}T$ coordinates, is defined in the following way. The intersection height of the principal ray with the first polar tangent plane is called Y_e . Then $S = Y_1 - Y_e$ and $T = H_1/l_{o1}$, (Fig. 2) and hence

$$\begin{aligned} \sigma &= l_{o1}/(l_{o1} - p), & \bar{\sigma} &= -p l_{o1}/(l_{o1} - p) \\ \tau &= -1/l_{o1}, & \bar{\tau} &= 1; & \bar{g} &= 1. \end{aligned}$$

Note that since S , T are coordinates defined at the first

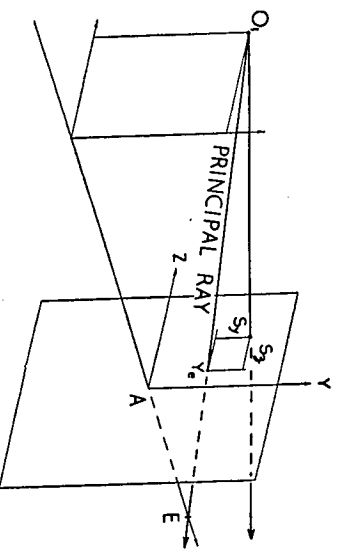


FIG. 2. Paracanonical ($\bar{O}T$) coordinates. E = center of entrance pupil, $T_y = H_y/l_{o1}$, $T_z = H_z/l_{o1}$.

surface only, they have no counterparts at any other surface, i.e., there are no variables S_j , T_j .

By using Eqs. M(12.61, 13.6) the expression for h_j becomes

$$h_j = N_1 T / N_j v_{aj}$$

where $T = V_1 / v_{b1} = \text{constant}$, V_1 referring to the principal ray.

Thus the expression for h_j is now the same for any position of the object and entrance pupil, their positions being contained in T .

The main advantage of paracanonial coordinates is that $t_{01} \equiv 0$ [see M, Sec. 13(e)] when the position of the object and entrance pupil planes are those used in the definition of the coordinates. Then expressions of the type

$$G_{\mu\nu j}^{(n)} = G_{\mu\nu p j}^{(n)} y_{01} + G_{\mu\nu q j}^{(n)} v_{01} \quad \text{M(8.8)}$$

become

$$G_{\mu\nu j}^{(n)} = G_{\mu\nu a j}^{(n)} s_{01}$$

and similarly

$$\Delta A_j = \Delta A_{aj} s_{01}.$$

Thus only the a coefficients are required to describe the aberrations associated with the various surfaces of an optical system.

As explained in M, Sec. 12(b), the change to paracanonial coordinates results in N_1/g appearing in equations where N_1 appeared previously. Since $g=1$ for $\bar{O}T$ coordinates, by making N_1 unity, N_1/g can be omitted from these equations. In the following sections of this paper N_1/g is taken as unity since this can always be arranged.

3. RAY TRACE THEORY

Consider Eqs. M(12.9) written in full. For this purpose, to simplify the printing, where s_y, \dots occur as subscripts, s and y will be written side by side.

Then

$$\begin{aligned} Y_j &= y_{aj}(S_y + \delta_{syj}) + y_{bj}(T_y + \delta_{tyj}) \\ Z_j &= y_{aj}(S_z + \delta_{szj}) + y_{bj}(T_z + \delta_{tzj}) \\ V_j &= v_{aj}(S_y + \delta_{syj}) + v_{bj}(T_y + \delta_{tyj}) \\ W_j &= v_{aj}(S_z + \delta_{szj}) + v_{bj}(T_z + \delta_{tzj}) \end{aligned} \quad (1)$$

where

$$\begin{aligned} \delta_{syj} &= - \sum_{i=1}^{j-1} \Delta A_{sybi} \\ \delta_{szj} &= - \sum_{i=1}^{j-1} \Delta A_{szbi} \\ \delta_{tyj} &= + \sum_{i=1}^{j-1} \Delta A_{tyai} \\ \delta_{tzj} &= + \sum_{i=1}^{j-1} \Delta A_{tza i} \end{aligned} \quad (2)$$

Writing δ_{yb} for δ_{sy} and ΔA_{yb} for $\Delta A_{syb} \dots$ we obtain

from (2) $\delta_{yb(j+1)} = \delta_{ybj} - \Delta A_{ybj}$, and analogously for quantities having the subscript " a ."

Hence

$$(S_y + \delta_{yb})_{j+1} = [(S_y + \delta_{yb}) - \Delta A_{yb}]_j$$

and similarly

$$\begin{aligned} (S_z + \delta_{zb})_{j+1} &= [(S_z + \delta_{zb}) - \Delta A_{zb}]_j \\ (T_y + \delta_{ya})_{j+1} &= [(T_y + \delta_{ya}) + \Delta A_{ya}]_j \\ (T_z + \delta_{za})_{j+1} &= [(T_z + \delta_{za}) + \Delta A_{za}]_j. \end{aligned} \quad (3)$$

These equations enable the transfer to be made from the j th to the $(j+1)$ th surface, the Y, V at each surface being obtained from Eq. (1). Note that $(S+\delta)_1 = S$ and that $(S+\delta)_j$ stands for $(S+\delta_j)$, the equations being written as they are in order to indicate that $(S+\delta_j)$ is evaluated at each surface for use in Eq. (1).

The ΔA 's are calculated as follows. For spherical surfaces only, from M(60.3),

$$\Delta A = \mathfrak{D}I.$$

That is

$$\begin{aligned} \Delta A_{ya} &= \mathfrak{D}aI_y, & \Delta A_{yb} &= \mathfrak{D}bI_y, \\ \Delta A_{za} &= \mathfrak{D}aI_z, & \Delta A_{zb} &= \mathfrak{D}bI_z, \end{aligned} \quad (4)$$

where

$$\mathfrak{D} = \Delta N [s_0(\mathfrak{S} - 1) - t_0' x \mathfrak{S}] \quad (5)$$

and

$$I = cY - V.$$

\mathfrak{S} is given by Eq. M(62.4), the coefficients of which M(62.5) contain a quantity \mathfrak{P} . Now from M(56.72)

$$\mathfrak{P} = 1 - cx$$

whence

$$x = (1 - \mathfrak{P})/c.$$

By substituting this expression for x in M(79.11) we get an equation in \mathfrak{P} , namely

$$(1 + \zeta)\mathfrak{P}^2 + 2(c\eta - \zeta)\mathfrak{P} + (c^2\xi - 2c\eta + \zeta - 1) = 0.$$

By letting

$$p_2 = 1 + \zeta = 1 + V^2 + W^2$$

$$p_1 = c\eta - \zeta = c(YV + ZW) - p_2 + 1$$

$$p_0 = c^2\xi - 2c\eta + \zeta - 1 = c^2(V^2 + Z^2) - 2p_1 - p_2$$

then

$$\mathfrak{P} = [\sqrt{(p_1^2 - p_0 p_2)} - p_1] / p_2$$

The positive root is always taken since $\mathfrak{P} \rightarrow 1$ in the paraxial limit (Y, Z, V, W infinitesimal).

When returning to \mathfrak{S} we consider, for convenience, a quantity

$$\bar{\mathfrak{S}} = (1 - k)\mathfrak{S},$$

M(62.4) then becomes

$$\bar{\mu}_2 \bar{\mathfrak{S}}^2 + 2\bar{\mu}_1 \bar{\mathfrak{S}} + \bar{\mu}_0 = 0$$

where

$$\bar{\mu}_2 = k^2 \mu_2 / (1 - k)^2, \quad \bar{\mu}_1 = k^2 \mu_1 / (1 - k), \quad \bar{\mu}_0 = k^2 \mu_0.$$

Thus

$$\bar{\mu}_0 = (k^2 - 1)(1 + \zeta) = (k^2 - 1)p_2$$

$$\bar{\mu}_1 = k^2(c\eta - \zeta) + (1 + \zeta)\mathfrak{P} = k^2p_1 + p_2\mathfrak{P}$$

$$\bar{\mu}_2 = k^2(c^2\xi - 2c\eta + \zeta) - (1 + \zeta)\mathfrak{P}^2 = k^2(p_0 + 1) - p_2\mathfrak{P}^2$$

since $\mathfrak{C} = 0$ for spherical surfaces [see M(56.71)].

Therefore

$$\bar{\mathfrak{E}} = [\pm\sqrt{(\bar{\mu}_1^2 - \bar{\mu}_0\bar{\mu}_2) - \bar{\mu}_1}]/\bar{\mu}_2$$

the positive root being taken when k is positive, the negative root when k is negative.

Equation (5) now becomes

$$\mathfrak{D} = N'[s_0(\bar{\mathfrak{E}} + k - 1) - t_0'x\bar{\mathfrak{E}}]$$

Let $\bar{\mathfrak{E}} + k - 1 = \bar{Q}$ say, then

$$\mathfrak{D}_a = N'[y_a\bar{Q} - v_a'x\bar{\mathfrak{E}}]$$

$$\mathfrak{D}_b = N'[y_b\bar{Q} - v_b'x\bar{\mathfrak{E}}]$$

since

$$\mathfrak{D} = \mathfrak{D}_a + \mathfrak{D}_b$$

and thus the $\Delta\Lambda$'s can be obtained from Eq. (4).

The displacements of the ray intersection in the ideal image plane from the ideal image point can be found from M(7.5). Thus, keeping in mind that $t_{01} = 0$,

$$\epsilon_{ky}' = \mu \sum_{i=1}^k \Delta\Lambda_{y ai}, \quad [\mu = 1/(N_k'v_{ak}')]]$$

$$\epsilon_{kz}' = \mu \sum_{i=1}^k \Delta\Lambda_{z ai}.$$

From Eqs. (2) and (3) it will be seen that

$$\sum_{i=1}^k \Delta\Lambda_{y ai} = (T_y + \delta_{ya})_k' - T_y$$

$$\sum_{i=1}^k \Delta\Lambda_{z ai} = (T_z + \delta_{za})_k' - T_z.$$

In practice, the object point is chosen to lie in the tangential plane, so that

$$\sum_{i=1}^k \Delta\Lambda_{z ai} = (T_z + \delta_{za})_k'.$$

As a check on the computation, it is convenient to calculate at each surface a quantity αE_x which is defined by M(4.3, 4.8) and is an optical invariant. Namely,

$$\begin{aligned} \alpha E_x &= \alpha N r I_x \\ &= \alpha N (ZV - YW) \\ &= N (ZV - YW) / \sqrt{(1 + V^2 + W^2)} \quad [\text{from M(3.2)}] \\ &= N (ZV - YW) / \sqrt{p_2}. \end{aligned}$$

It will be noted that this invariant is only useful in checking skew rays; for tangential rays it is necessarily zero.

As explained in M, Sec. 86, the refractive indices of an optical system are chosen appropriate to a particular wavelength (color) of light. These values of N determine a particular value of v_{ak}' and hence the position of the ideal image plane.

To trace rays of a different color, a new set of values of N are required which in turn fix the position of a new ideal image plane. Thus the scheme as presented so far gives the displacements of rays in the ideal image planes corresponding to the color of the rays traced. For chromatic aberrations, however, we are interested in the intersection of rays of all colors in a fixed image plane, the position of which is determined for some "base" color. In this case the displacement should be calculated from M(86.1), namely

$$\epsilon_{ky}' = {}_0l_{0k}'V' - Y' - {}_0h_k', \quad \epsilon_{kz}' = {}_0l_{0k}'W' - Z'$$

where

$${}_0l_{0k}' = {}_0y_{ak}' / {}_0v_{ak}', \quad {}_0h_k' = {}_0T_y / ({}_0N_k'{}_0v_{ak}').$$

The preceding subscript "0" refers to quantities associated with the base color.

For a new image plane F' which is at a distance x' from the ideal plane

$$\tilde{l}_{0k}' = {}_0l_{0k}' + x', \quad \tilde{h}_k' = {}_0h_k' + x'{}_0v_{bk}'{}_0T_y$$

and

$$\tilde{\epsilon}_{ky}' = \tilde{l}_{0k}'V' - Y' - \tilde{h}_k', \quad \tilde{\epsilon}_{kz}' = \tilde{l}_{0k}'W' - Z'.$$

4. SAMPLE CALCULATION

The system chosen as an example is the Pantoskop, Fig. 3, a six-surface, wide-angle lens system (Flügge,³ p. 283). The dimensions of this system have been adjusted to give it a focal length of unity ($f' = 1/v_{pk}'$). The diaphragm is 0.061849 unit behind the third surface. Table I shows the paraxial traces required to calculate p the position of the entrance pupil.

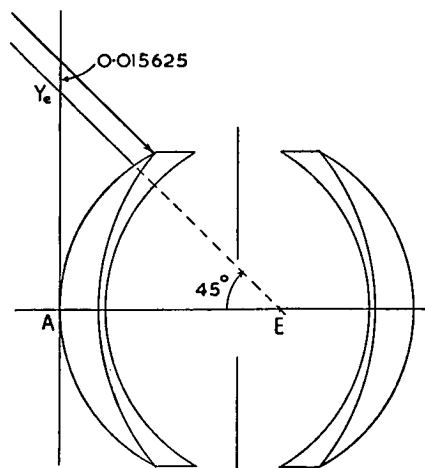


FIG. 3. Pantoskop wide-angle lens.

³ J. Flügge, *Das Photographische Objektiv* (Springer-Verlag, Berlin, Vienna, 1955).

TABLE I.

	$j \rightarrow$	1	2	3	4	5	6
	c	13.45565	9.648182	13.13822	-13.13822	-9.648182	-13.45565
	d		0.017357	0.001895	0.123697	0.001895	0.017357
	N	1.0	1.5331	1.6079	1.0	1.6079	1.5331
	k	0.652273	0.953480	1.6079	0.621929	1.048790	1.5331
	v_p	0.0	4.678890	4.873611	0.571964	-3.810739	-3.598413
p ray	$*y-dv=y_p$	1.0	0.918786	0.909549	0.838798	0.846021	0.908481
	$cy-v=v_p$	13.45565	4.185726	7.076241	-11.59228	-4.351828	-8.62579
	$(1-k)i+v=v_p'$	4.678890	4.873611	0.571964	-3.810739	-3.598413	1.0
	v_q	1.0	0.652273	0.616848	1.091591	1.411640	1.409813
q ray	y_q	0.0	-0.011322	-0.012491	-0.147518	-0.150194	-0.174664
	i_q	-1.0	-0.761508	-0.780957	0.846533	0.037455	0.940410
	v_q'	0.652273	0.616848	1.091591	1.411640	1.409813	0.908480
$d=0.061849, p = -(y_{q3}-v_{q3}'d)/(y_{p3}-v_{p3}'d)=0.091520$ [from M(33.1)]							

Suppose that the object point is at infinity and in the tangential plane. Then

$$\sigma=1, \bar{\sigma}=-0.091520, \tau=0, \bar{\tau}=1$$

Therefore the initial data for the a and b paraxial ray traces are;

$$a \text{ ray } y_1=1, v_1=0; b \text{ ray } y_1=0.091520, v_1=1$$

Suppose the principal ray from the object makes an angle of 45° with the axis of the system, hence $V_1=1$. The ray traced in Table II is the one of this pencil whose S_y, S_z are 0.015625, 0.015625. $T_y=V_1/v_{b1}=1$; $T_z=0$. Then

$$l_{0k}'=y_{ak}'/v_{ak}'=0.908481, h_k'=T_y/v_{ak}'=1.0$$

$$\epsilon_{ky}'=(T_y+\delta)_k'-T_y=l_{0k}'V'-Y'-h_k'=0.003189$$

$$\epsilon_{kz}'=(T_z+\delta)_k'=l_{0k}'W'-Z'=0.002148.$$

Suppose a new image plane is distant 0.015625 from the ideal image plane. Then

$$x'=0.015625, \bar{l}_{0k}'=0.924106, \bar{h}_k'=1.015625$$

$$\bar{\epsilon}_{ky}'=0.003498 \quad \bar{\epsilon}_{kz}'=0.002428$$

5. RAY TRACE PROGRAM

The scheme has been programed for a Deuce machine for systems containing 15 surfaces or less. When working to an accuracy of six decimal places, the upper limit of the machine capacity is 2^7 . This restricts the initial angles of a ray to be less than 80° . Note that the scheme will not work for angles $\geq 90^\circ$ in any case. Since S_y, S_z are the Cartesian coordinates of the intersection of a ray with the first polar tangent plane, with the intersection of the principal ray as origin, it is comparatively simple to program the machine automatically to trace a pencil of rays, given the principal ray data. If we remember that for an object point in the tangential plane $T_y=\text{constant}$ and $T_z=0$, the program works as follows.

Along with the normal lens parameters, the radii ρ of the rims of the components and the radius ρ_D and

position of the diaphragm are fed in as initial data. Following these are the initial data of the paraxial traces $y_{a1}, v_{a1}, y_{b1}, v_{b1}$, which, as mentioned earlier, contain the positions of the object and entrance pupils inherent in them. The machine then calculates the a and b paraxial coefficients of the system. It then calls for a "ray data card" on which is punched in binary V_1 of the principal ray, S_y, S_z , the "scan interval" $\bar{\delta}$ and x' . From V_1 and v_{b1} the machine calculates T_y, l_{0k}', h_k' which are stored for the duration of the trace. A pair of digits associated with V_1 are then examined to determine what type of trace is required, the four possible types being (1) full pencil, (2) tangential fan, (3) sagittal fan, (4) single ray. For the first three, S_y, S_z must be zero on the ray data card. After assuming that a full pencil is required, the machine takes as its starting data S_y, S_z from the card and hence traces the principal ray first. Now, at any curved surface $c^2(y^2+z^2)+\mathfrak{P}^2-1=0$ and thus only if $c^2\rho^2+\mathfrak{P}^2-1 \geq 0$ will the ray pass through the surface (for plane surfaces $\rho^2 \geq Y^2+Z^2$). At each surface this is checked and only if the ray is not vignettted will the trace continue. The coordinates of intersection of a ray in the plane of the diaphragm are found using Eq. M(3.3) and the ray is not vignettted here if $\rho_D^2 \geq \bar{y}^2+\bar{z}^2$. After tracing the principal ray, the values of $\epsilon_{ky}', \epsilon_{kz}', V_k', W_k'$ are stored. Then $\bar{\delta}$ is added to S_z and the trace reentered. Thus a sagittal fan is traced. This process continues until vignetting occurs somewhere in the system. Then S_z is made zero and $\bar{\delta}$ added to S_y and the trace restarted. By this means the first quadrant is scanned by a "grid of rays" spaced at intervals of $\bar{\delta}$. The first quadrant scan is terminated when a ray whose S_z is zero will not pass through the system. Then the fourth quadrant is scanned in a similar manner by subtracting $\bar{\delta}$ from S_y each time a ray is vignettted. When the lower quadrant has been filled, all the results are punched out. Along with the ϵ 's and V', W' for each ray, each card contains the Cartesian coordinates of the particular ray, taking $\bar{\delta}$ as the unit of length. If, in the example $\bar{\delta}=2^{-8}$, the (y,z) of the ray would be $(+4, +4)$. If an x' is punched on the ray data

TABLE II.

$j \rightarrow$	1	2	3	4	5	6	k'
c	13.45565	9.648182	13.13822	-13.13822	-9.648182	-13.45565	
d		0.017357	0.001895	0.123697	0.001895	0.017357	
N'	1.5331	1.6079	1.0	1.6079	1.5331	1.0	
k	0.652273	0.953480	1.60790	0.621929	1.048790	1.5331	
y_a	l_1	1.0	0.918786	0.909549	0.838798	0.846021	0.908481
v_a	l_2	0.0	4.678890	4.873611	0.571964	-3.810739	-3.598413
y_b	l_3	0.091520	0.072766	0.070751	-0.070751	-0.072766	-0.091520
v_b	l_4	1.0	1.080486	1.062881	1.143938	1.062881	1.080486
$*(t_5 - t_{29}) = (S_y + \delta)$	l_5	0.015625	0.018790	0.018872	0.018930	0.016044	0.016608
$*(t_6 - t_{30}) = (S_z + \delta)$	l_6	0.015625	0.017131	0.017117	0.016438	0.015211	0.015284
$*(t_7 + t_{27}) = (T_y + \delta)$	l_7	1.0	1.084969	1.075404	1.078590	0.993350	1.005195
$*(t_8 + t_{28}) = (T_z + \delta)$	l_8	0.0	0.040444	0.042069	0.004702	-0.031529	-0.030013
$l_{15} + l_{27} = Y$	l_9	0.107145	0.096212	0.093251	-0.060433	-0.058708	-0.076908
$l_{16} + l_{28} = Z$	l_{10}	0.015625	0.018683	0.018546	0.013456	0.015163	0.016632
$l_{25} + l_{47} = V$	l_{11}	1.0	1.260209	1.235003	1.244667	0.994674	1.026336
$l_{26} + l_{48} = W$	l_{12}	0.0	0.123855	0.128138	0.014780	-0.091479	-0.087426
$l_{11}^2 + l_{12}^2 + 1 = p_2$	l_{13}	2.0	2.603468	2.541651	2.549415	1.997744	2.061009
$c(l_9 l_{11} + l_{10} l_{12}) - l_{13} + 1 = p_1$	l_{14}	0.441708	-0.411322	0.002641	-0.563787	-0.420950	0.020651
$c^2(l_9^2 + l_{10}^2) - 2l_{14} - l_{13} = p_0$	l_{15}	-0.760691	-0.886638	-0.986559	-0.760184	-0.813600	-0.981332
$l_{14}^2 - l_{13} l_{15} = \mathfrak{P}$	l_{16}	1.716488	2.477519	2.507495	2.255881	1.802564	2.022961
$(\sqrt{l_{16} - l_{14}})/l_{13} = \mathfrak{P}$	l_{17}	0.434220	0.762573	0.621984	0.810282	0.882769	0.680083
$(k^2 - 1)l_{13} = \bar{\mu}_0$	l_{18}	-1.149079	-0.236594	4.029387	-1.563311	0.199696	2.783178
$k^2 l_{14} + l_{17} l_{18} = \bar{\mu}_1$	l_{19}	1.056369	1.611393	1.587694	1.847675	1.300517	1.450195
$k^2(l_{15} + 1) - l_{17}^2 l_{13} = \bar{\mu}_2$	l_{20}	-0.275278	-1.410904	-0.948523	-1.581077	-1.351770	-0.909367
$l_{16}^2 - l_{18} l_{20} = \bar{\mathfrak{C}}$	l_{21}	0.799600	2.262777	6.342740	0.942187	1.961288	4.633997
$(\sqrt{l_{21} - l_{19}})/l_{20} = \bar{\mathfrak{C}}$	l_{22}	0.589098	0.075937	-0.981300	0.554693	-0.073935	-0.772490
$(1 - l_{17})l_{22}/c = x\bar{\mathfrak{C}}$	l_{23}	0.024770	0.001869	-0.028234	-0.008010	0.000898	0.018366
$k + l_{22} - 1 = \bar{Q}$	l_{24}	0.241372	0.029417	-0.373400	0.176622	-0.025145	-0.239390
$[l_{17} l_{24} - (l_2^*)l_{23}]N' = \mathfrak{D}_a$	l_{25}	0.192364	0.028815	-0.323476	0.189132	-0.027657	-0.235847
$[l_3 l_{24} - (l_4^*)l_{23}]N' = \mathfrak{D}_b$	l_{26}	-0.007165	0.000248	0.005880	-0.006404	0.001317	0.003543
$(cl_9 - l_{11})l_{25} = \Delta\Lambda_{ya}$	l_{27}	0.084969	-0.009565	0.003185	-0.085239	0.011844	-0.002006
$(cl_{10} - l_{12})l_{25} = \Delta\Lambda_{za}$	l_{28}	0.040444	0.001625	-0.037367	-0.036231	0.001516	0.032161
$(cl_9 - l_{11})l_{26} = \Delta\Lambda_{yb}$	l_{29}	-0.003165	-0.000082	-0.000058	0.002886	-0.000564	0.000030
$(cl_{10} - l_{12})l_{26} = \Delta\Lambda_{zb}$	l_{30}	-0.001506	0.000014	0.000679	0.001227	-0.000072	-0.000483
$N(l_{10} l_{11} - l_9 l_{12})$	l_{31}	0.015625	0.017827	0.017614	0.017641	0.015616	0.015862
$l_{31}/\sqrt{l_{13}} = \alpha E_x$		0.011049	0.011049	0.011049	0.011049	0.011049	0.011049

card, the final results punched are those for the new image plane position. After the results are obtained, a new ray data card is called in and the program restarts.

The tangential and sagittal fans are variations of the full scan. Incidentally, if $V_1 = 0$ only the first quadrant is scanned. Apart from specifying the location of the ray, the (y, z) in the case of the full pencil give the apparent shape of the entrance pupil. The V_k', W_k' are punched out in case the designer wishes to find x_s, x_T (M 203.51, 53).

Since the program is designed to handle chromatic effects, the ϵ 's are calculated from the equations involving l_{0k}' . The l_{0k}', o_{hk}' are stored in the machine and when a new set of refractive indices are fed in corresponding to a new color, the machine is told to use the stored l_{0k}' in the calculation of the ϵ 's, the o_{hk}' being selected from the store to correspond to the V_1 on the ray data cards for the new color. This of course presupposes that at least one ray in the base color has

been traced for each of the V_1 required in other colors.

Note that at each surface, $\mathfrak{P} \leq 1$, $|\bar{\mathfrak{C}}| \geq |1 - k|$, l_{16} must be positive, and l_{21} must be positive. If l_{16} is negative, the ray has missed the surface ($y^2 + z^2 > r^2$) and if l_{21} is negative, total internal reflection has occurred. Thus the computation can be checked at each surface, the calculation of αE_x checking the ray trace through the entire system.

ACKNOWLEDGMENTS

I wish to thank Dr. H. A. Buchdahl for his assistance in the design of the program and Mrs. Brown, who checked the coding and cards as well as computing the necessary check ray traces. The checking of the program on the Deuce was made much easier by the generous help given by Mr. R. Smart and his staff associated with the computer at the University of New South Wales.